

Task force turns on sponsor (DoE)

The US national laboratories are a large part of the richest country's resources in research, and are being wasted by mismanagement by the Department of Energy — according to a DoE task force now reporting.

THE US Department of Energy (DoE) has only itself to blame. A year ago, it commissioned a powerful task force to suggest what should be done with the ten national laboratories in its control, most of them originally founded during the Manhattan Project and many of them bereft of a pointed role for several years. The task force has now reported (see page 463) in brisk and incisive language. The main part of its document is an entertaining and instructive read. So does that mean that the national laboratories are now destined for oblivion? On the contrary, from Argonne in Illinois and Oak Ridge in Tennessee to Livermore in California, the laboratories can relax. In forthright language, the task force has indicted not the laboratories but their sponsor, the Department of Energy.

It is an extraordinary and instructive tale, well illustrated by what the task force says about the management by DoE of the programme to rid former nuclear weapons sites of radioactive contamination. Nobody knows for sure what the eventual cost will be, but some estimates put the total at \$1,000 billion. The national laboratories, which played a part in creating the means by which the radioactive waste was created in the first place, have also been deeply engaged in the difficult assessment of the scale of the problem, in investigating and developing technical means of ridding soil and ground-water of contamination and so on. But the management of the programme has been firmly in the hands of DoE, which has bungled it.

Credibility

The task force agrees with DoE's critics that there is precious little to show, so far, for the billions already spent on clean-up, that DoE's standard response to criticisms of its management of the programme has been to hire more managers and that, in the absence of a considered plan of what might be achieved by when, it has been (expensively for the public purse) hostage to claims of damages for non-compliance with regulations promulgated by other departments than itself. The task force notes that DoE's credibility is poor in the literal sense that people in general no longer believe what it says about the clean-up programme. The task force scathingly reminds DoE of its duty to carry through the programme professionally, and in the interests of the public. Its remedy: that there should be an independent committee to manage the programme and that one of the national laboratories should be designated to take the lead.

DoE's recent management of its laboratories is similarly

found to have been deficient to the point of fecklessness. As the weapons programme has declined, laboratories originally established for particular roles have been encouraged to diversify — which in practice has meant competing for other research funds for other kinds of projects. There is a fable in Aesop of a dog carrying a bone in its mouth when crossing a river; seeing the reflection in the water, the dog drops its bone to grab for the reflection. That, the task force suggests, is how the national laboratories have been compelled by DoE to behave in the competition for novel funds, skimping on existing programmes so as to compete for the new and trendy. Except that there is not one dog, but close on a dozen.

Bureaucracy

The bureaucracy is something else again. Tales of over-detailed instructions from head office are so familiar in national laboratories that they have ceased to be remarkable, but it is really necessary that DoE should have spent \$500,000 on a document (most of it salary cost, no doubt) to show cause why a radioactive storage tank at Hanford in Washington state should not be decontaminated just yet? To be fair, DoE has been sitting uncomfortably beneath the eye of Congress ever since it inherited research and development in nuclear energy from the old Atomic Energy Commission, and there is nothing like a congressional edict to start the paperwork flowing. But that does not excuse the exquisitely pointless procedures on which the department's civil servants lavish their time — and other people's. And it does not wholly account for the most chilling reference in the report to the sense of intimidation within the department's staff born of the fear that they "might go to jail". Curiously, the department's leaders have not been conspicuous by their protests at this state of affairs in the past few years.

What is to be done? The task force's chief and daring recommendation is that the national laboratories should be differentiated from each other by goal and that they should then be "corporatized", which is a word for a novel form of public ownership not yet attempted. The idea is that the laboratories (possibly excepting those with a continuing responsibility for nuclear weapons design) should be owned in the public interest by trustees, managed as if they were autonomous research corporations and financed by government and other agencies seeking to buy research services. It is to be hoped that this adventurous idea will be given the sympathetic study it deserves. The obvious danger is that if



the US national laboratories are allowed to continue to drift as in the past several years, the quality of those who work in them will disastrously decline. Out of politeness, no doubt, the task force has evidently shrunk from saying that no such innovation will be given a chance if the old guard at the DoE has its say.

The task force, meanwhile, has done as much as it could to improve morale at the laboratories. Its attack on the bureaucracy rather than the laboratories themselves will no doubt evoke a chorus of "We've been saying that for years!" True, there is some talk of "downsizing" (shamefully accompanied by the politically correct "rightsizing" in a few references), but that should not worry the good people at these excellent laboratories. For them, the task force has a cheerful prospectus. Why not take energy research and policy more seriously? Why not lump energy conservation with the conservation (or economical use) of other industrial raw materials in a new attack on "industrial ecology"? And then there are always the clean-up programme and the fight against the proliferation of nuclear weapons. It is an almost thrilling tale. DoE may be too jaded to listen to it. Will the new Congress simply pocket as much as it can of the \$6 billion a year the national laboratories spend, and not bother to listen? That is what will matter. □

Intellectual wars

The US dispute with China over intellectual property rights threatens lasting and avoidable trouble.

THE US dispute with China over the protection of intellectual property is a high-grade dispute and could quickly escalate if it is not defused. The US Trade Representative, Mr Mickey Kantor, has decreed that roughly \$1 billion's worth of Chinese exports to the United States will be subjected to 100 per cent tariffs from the end of the month, and will thus be unaffordable. The cause of the trade war is China's unwillingness to protect US intellectual property with the diligence the United States requires of its trading partners. It is a pity that Kantor has not found safer ground on which to fight. China's unpredictable promise to retaliate will bring no good either.

Copyright is, of course, a crucial element in international trade. Devised to protect authors' rights in works their ingenuity has made possible, copyright is the means by which the authors earn a just reward and their emulators are spurred to do even better. But there is copyright and copyright. Even in the rich countries of the world, it is increasingly a problem, for the would-be legal purchasers of copyrighted software, to buy what they want without having to buy (at extra cost) extra bells and whistles for which they have no need. In the poorer countries of the world, the terms on which copyrighted materials are made available are crucial. In China, which is economically powerful in aggregate but still poverty-stricken per capita, Western royalties on software and compact discs would ensure no sales.

Kantor would be entirely within his right to indignation with China if there were grounds for believing that the

government condones the export from China of counterfeited software and the like. A sufficiently public display of such indignation would probably bring such practices to a halt. But the rest of his complaint would carry greater weight if he were able to show how the manufacturers of the US materials allegedly pirated in China had offered reasonable terms for their local reproduction, and had been refused. That vital missing element in the case should be plugged before the inevitable escalation sets in. □

Is plagiarism OK?

A bruising civil suit between two academics at the University of Hong Kong will do the university no good.

THE University of Hong Kong has had more on its mind these past two years than the prospect that it will be within the borders of China in less than 30 months. At the end of 1993, Hong Kong's appeal court confirmed an earlier court decision in favour of Dr Linda Koo, who had brought a civil suit against a close colleague, Dr Lam Tai Ying, for damages on the grounds of plagiarism. Koo and Lam are respectively a lecturer and a reader in the department of community medicine at the university. Koo's complaint was that Lam had used a questionnaire she had devised to inquire into the surprisingly high incidence of lung cancer among women in Hong Kong to devise his own questionnaire, to be used in a study of the surprisingly high incidence of lung cancer among women in Hong Kong.... A civil suit does not, of course, lead to a criminal trial. The questions to decide are not 'guilt' or 'innocence', but 'liability' and 'how much?'

The first thing to say about such a suit is that it should never happen (or have happened). However dark may be the thoughts that academics harbour against their colleagues "in the still night, when only the Moon rages", in the light of day a minimal sense of perspective would dictate that there must be better ways. That is all the more the case when the disputing academics belong not merely to the same university, but to the same department. But the University of Hong Kong seems not to have found the trick to head off either the original dispute or its inevitable successor, the formal question whether Lam should be dismissed for scientific misconduct. So a committee took evidence from all in sight (except for Koo, whom it failed to accommodate in its timetable) and decided that Lam should remain in post.

In Hong Kong and even elsewhere, it will seem a daring and even a foolhardy decision by a university (however properly arrived at) to fly in the face of a court decision, even in a civil suit. But there will also be many, in Hong Kong and elsewhere, who will be asking whether it can make sense that academics can take each other to court. Other questions, not yet answered, are why the same department of community medicine thought it a sensible use of resources to mount two such parallel investigations of the same problem that plagiarism could be judged beneficial. Squabbles like this, in short, give universities a bad name. Let us hope that Beijing will not have noticed. □

Cuts force UK astronomers to drop key role in gamma-ray satellite

London. British astronomers have been told that they will not be provided with funds they need to play a leading part in the construction and exploitation of a new ECU322-million (US\$396-million) gamma-ray satellite, *Integral*, initially proposed by Britain, which is planned for launch by the European Space Agency (ESA).

The lack of funds is a direct result of decisions by the Office of Science and Technology to put pressure on ESA to reduce its space science budget, and not to provide the Particle Physics and Astronomy Research Council (PPARC) with funding for any new domestic scientific projects in the financial year 1995–96.

Anticipating such a squeeze on its funding, PPARC had already asked ESA to delay any decision about allocating contracts for the payload for *Integral*, adopted in 1993 as ESA's next 'medium-scale' mission, until after a meeting of the space ministers of the agency's member states, planned for next November.

Britain is hoping to use this meeting to persuade other ESA members to accept a reduction in the agency's science budget. It argues that tactics that have already proved successful in reducing the overhead costs of the European Laboratory for Particle Physics (CERN) can also be applied to the space agency (see *Nature* 372, 712; 1994).

But Britain's request has been angrily rejected by Roger Bonnet, director of ESA's scientific programme. He says that, even if Britain is unable to contribute to the payload costs of *Integral* (the satellite itself will come out of ESA's central funds), the mission will still proceed, but on a reduced basis. Italy has already offered to take over some of Britain's planned role.

Two main instruments planned for *Integral* will be affected. The first is one of its two large gamma-ray telescopes, an caesium

iodide imager, for which a British research team headed by Tony Dean at the University of Southampton had already been identified as the most likely principal investigator, as much of the development work has been carried out in the United Kingdom.

Originally planned as a joint UK/Italian project, it now appears that the Italians may now be prepared to cover all the costs of the imager — and that, as a result, Italian scientists will have privileged access to data sent back over the first year, as is customary practice.

The second instrument is an optical transient camera, which had been planned as a joint project between Britain and Ireland. According to PPARC officials, it now looks as if this will be scrapped altogether.

David Hunt, the science minister, announced last week that PPARC has been awarded an overall budget increase of 4.8 per cent next year (see page 464). But most of this will go towards paying the extra costs of the UK subscription to CERN caused by the fall in the value of the British pound compared to the Swiss franc (in which currency the CERN subscription is paid).

There will also be an increase in the subscription to ESA, in line with previous commitments to the agency. But there will be no extra money left for increased spending on domestic programmes or new scientific starts.

Ken Pounds, PPARC's chief executive (and also professor of radioastronomy at the University of Leicester), says that part of the



Pounds: problems are 'price of success'.

problem has ironically been the success of British scientists in winning a substantial role in a number of existing ESA programmes, such as *Soho/Cluster* solar mission, each of which has required substantial funding.

"One could say that we are suffering from indigestion, having just eaten a large meal," says Pounds. "But the result is that we are going to miss the next meal, with pressure on the council's domestic budget likely to lead to British groups not being able to play a significant part in *Integral*."

British astronomers, angry that the scientific fruits of many years' work designing experiments for *Integral* are likely to be plucked from under their nose by others, see it differently.

They argue that are being denied funding because *Integral* has become a pawn in the government's efforts to reduce the size of its ESA subscription.

"*Integral* has become part of a negotiating ploy," complains Gerry Skinner of the University of Birmingham's school of physics and space research. "PPARC's decision not to provide any money next year means that Britain will lose out both by not being able to carry out the hardware development programme for the payload, and through being denied access to that fraction of the data which will be 'top-sliced' by those who provide the payload."

Pounds and Bonnet were due to meet earlier this week to discuss the impact of Britain's decision on ESA's plans. Despite Britain's request for a delay in decisions over the contractors for *Integral*'s payload, due in June, Bonnet is said to be keen to keep the mission on schedule in order to turn the agency's attention to its future programme, *Horizon 2000+*.

Even if Bonnet insists on proceeding with *Integral* without Britain, the agency's difficulties may be far from over. No one doubts the genuine desire (and ability) of Italian scientists to pick up the pieces. There is also an economic logic, given that the Italian aerospace company Alenia has already been selected as prime contractor for *Integral*.

A spokesman for the Italian Space Agency (ASI) said this week that various government agencies are discussing what level of contribution they may be able to make. But with the agency currently in turmoil, and facing a major task in renegotiating its previous commitments to the European agency, ESA officials are said to be cautious about how far Italian promises to fill the gaps left by Britain are likely to materialize.

David Dickson

Advisers ignored over observatory head

Munich. Italian astronomers are hoping to establish a better relationship with their new research minister, Giorgio Salvini, than they experienced with his predecessor Stefano Podestà. Last month, Podestà rejected without any explanation the nominee of the country's 15-strong council for astronomical research for the directorship of one of the most scientifically important of Italy's 13 observatories, the Brera Astronomical Observatory near Milan.

Instead, he renewed the term of the current director, Guido Chincarini, for a further three years. The move has angered astronomers, as it is the first time that the advice of

the council — an elected body of scientists responsible for advising the research minister on matters relating to astronomy — has been ignored.

Giancarlo Setti, the vice-president of the committee, says that the committee is not complaining about the scientific merits of the two candidates, both of whom, he says, are "excellent scientists". But he says that the committee is concerned about the way in which the former research minister — who was often criticized for failing to listen to the scientific community during his seven-month term of office — appears to have rejected the committee's advice. **A. A.**

Clinton puts freeze on US science budget

Washington. The Clinton administration presented a science and technology budget proposal to Congress on Monday which seeks to boost basic research in the universities and pursue technology programmes favoured by the president, while pursuing both goals within tight fiscal constraints.

The White House proposal for the 1996 fiscal year — which begins on October 1 — seeks to transfer \$1 billion from military to civil research programmes within an overall science and technology package that is frozen at this year's level of \$73 billion. Once inflation is taken into account, the proposals would result in a real cut of about 3 per cent in US government spending on research and development.

But the package is widely seen as less significant than Clinton's budget proposals for the previous two years, as the new Republican majorities in both houses of Congress are likely to use their power to change it almost beyond recognition.

Clinton announced no specific programme cuts in the science budget apart from the cancellation of the proposed Advanced Neutron Source (see right).

The administration stood by its plans significantly to increase spending in 1996 on selected technology programmes, including the Advanced Technology Programme and the Partnership for a New Generation of Vehicles (PNGV). "This is no time to waver on our commitment to these programmes," says Jack Gibbons, the President's science advisor.

Despite the freeze in the science and technology budget, no real attempt has been made to reach common ground with Congress by trimming programmes which Congress is preparing to oppose. That approach

reflected the overall budget proposal, which avoided specifying budget cuts and, critics say, would allow the US budget deficit to grow further. Many observers say that President Clinton, having won no political credit for last year's deficit reduction, declined to repeat the error. He intends to stick with his programmes, and to ensure that, if painful cuts come this year, the Republican Congress takes the blame.

"What you see is a re-ordering of priorities," Gibbons said on Monday. "This is a budget that emphasizes the basics of science and technology." By slowing down invest-



Gibbons: need to face 'tough times'.

ment in facilities, as well as in military research and development, the administration says it will be able to expand spending on all basic research by 3.5 per cent to \$14.5 billion, and on university research by 7 per cent, to \$12.5 billion.

Gibbons said that key areas of spending in the budget included more money for multi-agency partnerships with industry, such as the PNGV, more health and environmental research, and a leaner approach to space exploration at NASA.

Conceding that, once inflation is taken into account, the frozen overall budget figure represents a real budget cut, Gibbons said it was important to remember that the administration was facing "tough times", and that no part of the budget was going to escape scrutiny. "Some programmes are up," he said. "This proposal puts a bottom line on the importance which the president puts in these investments." **Colin Macilwain**

DoE: neutron source to be axed

Washington. The US Department of Energy announced on Monday that it is terminating work on the Advanced Neutron Source (ANS), planned to have been built at Oak Ridge National Laboratory, Tennessee. But it but failed to specify any of \$1.2 billion in cuts in applied research which it has promised to make over the next five years.

The department also announced a new, \$100 million programme — known as the Scientific Facilities Initiative — aimed at improving the access of university and industrial researchers to a wide range of synchrotrons and other facilities.

The new initiative is planned to come

from an increase in the department's spending on "fundamental science" from \$2.7 billion to \$2.8 billion.

The budget published by DoE said that half of this increase — \$50 million — is to be saved through unspecified 'Galvin task force reductions'. But the task force, which reported last week on the future of the energy department laboratories (see page 463), did not suggest any ways of saving money at the laboratories in the short term. "I don't know where that number comes from," Martha Krebs, director of the Office of Energy Research at DoE, said in Washington on Monday.

The DoE budget proposes a substantial expansion in spending on high energy physics, from \$642 million to \$686 million, as well as an increase in basic energy sciences from \$734 million to \$811 million. Most of both these increases would come from money that is being allocated to the Scientific Facilities Initiative.

Hazel O'Leary, the energy secretary, said that the increases "would put high energy physics back on course" after the cancellation of the superconducting supercollider in 1993, and "help to get industry and the universities into our tremendously valuable facilities."

Despite intensifying calls for the abolition of the Department of Energy, the administration plans to increase the department's total budget next year from \$17.5 billion to \$17.8 billion. O'Leary — standing in front of a cheque for \$14.1 billion made out to the US taxpayer — said that none of the savings that have been promised can be delivered in FY96. Congress is likely to view matters differently. **C. M.**

NSF: Research grants get 8 per cent boost

Washington. The National Science Foundation (NSF), the main source of non-bio-medical university research in the United States, wants to increase such funding by 7.6 per cent next year, despite tight budget constraints.

The Clinton administration is proposing that the total NSF budget should grow by 3 per cent to \$3.36 billion in 1996 — the lowest rate proposed by any administration in recent years. But by cutting back on education and the construction of research facilities, the agency proposes to boost its research grant funding by \$170 million to \$2.45 billion.

The budget proposal reflects the usual pattern of NSF's relationship with Congress: the agency tries to focus on its central task of awarding research grants, and Congress then insists on distributing more of its money to build scientific facilities, buy research equipment and boost education.

Last year, for example, Congress approved \$250 million for NSF facilities; the budget proposal rescinds half of that, and proposes just \$100 million for facilities next year. The administration is also proposing to make a sharp cut in funding for major research equipment — reducing this from \$126 million to \$70 million — and a 2 per cent cut in the NSF's \$600 million science education budget.

The boost to research funding is distributed evenly across disciplines, although the mathematical and physical sciences (up 8.3 per cent) and social sciences (up 8.0 per cent) do slightly better than others. Most of the "strategic" research areas favoured by the previous Democrat-controlled Congress are listed for a smaller-than-average increase. But the foundation's director, Neal Lane, says that almost two-thirds of NSF-supported research is now defined as "strategic." **C. M.**

NIH: in line with expectations

Washington. The 4.1 per cent budget increase being requested by President Clinton for the National Institutes of Health (NIH) in 1996 represents an increase that would give the NIH about \$11.8 billion to spend next year. Increases for individual institutes ranging from a little below 3 per cent for some, to nearly 5.5 per cent for the Office of AIDS Research.

Overall, the NIH's proposed budget increase is a little below the Department of Commerce's predicted increase (4.3 per cent) in the cost of doing biomedical research in 1996. "It's disappointing," says Sam Silverstein, president of the Federation of American Societies for Experimental Biology (FASEB).

But the budget request is larger than the general rate of inflation, and roughly the amount that most observers expected. Most of the money that the NIH spends on research will, as usual, be awarded to unsolicited grants selected through peer review. In addition — and providing that the new Congress raises no objections — extra money will as last year also be directed into research on HIV and AIDS, and towards improving the health of women and minorities.

A new area for targeted research, championed by Harold Varmus, the director of the NIH, will be preventive research in

brain disorders. Other targeted areas for funding will be environmental cancer and gene therapy research.

Although observers say that are not surprised by the president's request, it falls far short of the 15 per cent increase that internal staff at the NIH estimate is needed to fund one third of grant proposals — at present, only 25 per cent are funded — and to provide the necessary buildings and equipment. But the NIH was not asking for this figure.

Such figures go some way to explaining a call by FASEB last year to increase the NIH's budget by 10 per cent. A staff member says that this percentage was based on an assessment of the need to increase the percentage of research proposals being funded, to expand training and to provide more funds for instrumentation.

Given this call from FASEB, Silverstein's view that the current budget is a missed opportunity for the NIH is not entirely surprising. "At a time when the president is saying the country should invest, this budget is not an investment for the NIH, but holding the line," says Silverstein. For many, however, the budget was as much as could be expected in the circumstances — and FASEB has several months in which to lobby Congress to increase the president's request.

Helen Gavaghan

NASA: new starts despite tough year

Washington. Although its budget request for next year of \$14.36 billion is \$200 million less than this year's allocation, the National Aeronautics and Space Administration (NASA) intends to fund modest new initiatives in space science and technology, as well as a new reusable rocket programme (see *Nature* 373, 180; 1995).

The agency is seeking \$6 billion for its science, aeronautics and technology account, about the same level as in the current year. Funding for the space shuttle and space station programmes remains virtually constant, at \$5.5 billion.

The agency's budget request includes money that would allow it to begin work on two infrared astronomy missions that have long been on the drawing board: \$15 million for the Space Infrared Telescope Facility (SIRTF), and \$49 million for the Stratospheric Observatory for Infrared Astronomy (SOFIA). It also includes \$50 million to develop a second spacecraft in the Mars Surveyor series, to be launched in 1998. The potential contractors for the Mars orbiter/lander will be chosen this spring.

The Discovery line of small planetary missions is requesting an additional \$36 million to begin work on at least one new mission to follow the Mars Pathfinder and Near Earth Asteroid Rendezvous (NEAR) missions already in development. NASA

expects to make a selection from a list of candidate missions later this month.

In addition to \$159 million for the Reusable Launch Vehicle programme, the agency is also asking for \$30 million to kick off what Administrator Daniel Goldin calls a "revolutionary" new technology development effort, called New Millennium. This aims at reducing both the cost and weight of current science spacecraft by a factor of ten.

The first demonstration mission, which would probably fly a miniaturized science payload to a nearby target such as an asteroid, could be launched as early as 1997. Goldin expects New Millennium to run as a continuing technology development effort, costing approximately \$50 million a year.

Funding for these new initiatives is being offset by general savings in operations, and by the reduced budgetary needs of the Cassini Saturn mission as it approaches its 1997 launch. There is, however, a wild card.

Money for the proposed Gravity-Probe B mission, designed to test Einstein's Theory of Relativity, is not included in NASA's 1996 budget request, as this awaits a recommendation by the National Academy of Sciences this summer on whether to proceed with the project. If the answer is yes, NASA will have to find \$51 million in a tight budget to pay for it next year.

Tony Reichhardt

Other highlights . . .

Department of Interior: The 1996 request is for \$7.564 billion, \$215 million higher than last year. The administration is proposing slight raises for both the US Geological Survey (\$586 million, up \$15 million from 1995) and the National Biological Service (\$173 million, up \$6 million). The implication is that if either of these agencies are to be killed this year, Congress will have to do it.

The USGS request includes additional money for digital mapping, earthquake preparedness, seafloor mapping, and data collection and reporting under the National Water Quality Assessment programme. The survey plans to save \$4.6 million by closing 54 water resources research institutes around the country.

Environmental Protection Agency: EPA is seeking \$7.4 billion in next year's budget, \$138 million more than it received in 1995. The agency's Office of Research and Development wants an increase of \$84 million on last year, increasing its budget to \$630 million — although some of that rise is due to research funds being reclassified from elsewhere in the agency.

The agency is also asking Congress to approve spending an extra \$42 million on its recently expanded programme of external research grants, and \$5 million for more graduate student fellowships.

Department of Defense: Cuts of almost \$1 billion are being proposed from the \$35 billion which the Department of Defense spends on the development and test of military equipment, and which the US government classifies as "research and development." But the department wants to spend \$1.2 billion — the same as last year — on its university research, even though this lacks direct military application, and has come under attack from defence hawks in Congress.

Department of Commerce: The National Institute of Standards and Technology (NIST) continues to pursue the rapid expansion of its industrial technology programmes (see *Nature* 373, 374; 1995). The NIST budget would rise from \$854 million to \$1.023 billion, with the Advanced Technology Programme (ATP) growing from \$430 million to \$490 million.

This is half the growth rate which the Clinton administration had previously advocated, but is still unlikely to appease Republican opponents of programme. The budget also proposes significant expansions in the Manufacturing Extension Partnership, and in NIST's intramural research programme. **C. M. & T. R.**

Prospects dim for agreement on new climate targets

London. The chances of reaching international agreement in the near future on binding targets for a long-term reduction in greenhouse gas emissions are rapidly diminishing. Few now expect the countries that signed the United Nations Framework Convention on Climate Change (FCCC) in Rio de Janeiro in 1992, which hold their first full post-Rio meeting in Berlin at the end of March, will agree such a goal.

As delegates gathered in New York last Monday (6 February) for their final preparatory meeting — referred to as Intergovernmental Negotiating Committee (INC) 11 — the most that supporters of binding reduction targets were expecting was a pledge to begin negotiations on a post-2000 protocol. Developed countries may also agree to adopt an interim goal of stabilizing CO₂ emissions after 2000 at 1990 levels.

The commitment enshrined in the treaty signed at the 'Earth Summit' was that such countries should aim to return their greenhouse gas emissions to 1990 levels by the year 2000. This was to be the first step towards the ultimate objective of stabilizing greenhouse gas concentrations in the atmosphere at a sufficiently low level to prevent dangerous interference with the climate system (see *Nature* 371, 274; 1994).

But the measures introduced by most countries so far seem to be falling short of their targets. According to an analysis by the interim secretariat of the FCCC of the 15 country programmes received so far, only five countries are expected to return to 1990 CO₂ levels or less by the year 2000 without additional measures.

Nevertheless, many countries are confident that they will meet the targets. Rafe Pomerance, deputy assistant secretary for environment and development in the bureau of oceans and international environmental and scientific affairs in the US State Department, says that voluntary measures in the US programme are only just starting to take off. "The President has made a commitment to return to 1990 levels by the year 2000 so that's our national commitment."

A handful of European Union (EU) countries are likely to achieve the target among them the United Kingdom, but the EU as a whole seems unlikely to do so. Nevertheless, the 12 environment ministers of the EU have joined other countries in stating that the current commitment is inadequate to achieve the main objective and must be strengthened.

This so-called review of the "adequacy of commitment" is likely to be one of the more contentious issues discussed at Berlin, particularly as the Intergovernmental Panel of Climate Change (IPCC)'s Special Report

was officially launched on the first day of INC 11. The report reiterates earlier conclusions by the IPCC that stabilization of greenhouse gas emissions at 1990 levels will not stabilize atmospheric concentrations at non-dangerous levels.

But the EU ministers drew back from giving their full support to a proposal for a draft protocol, already tabled by Trinidad and Tobago on behalf of the Alliance of Small Island States, which commits signatories to the so-called 'Toronto Target'. This specifies a goal of at least 20 per cent reductions of anthropogenic emissions of CO₂ compared to 1990 levels by the year 2005.

Instead, the ministers decided to seek agreement for stabilization of CO₂ emissions and to begin negotiations on how the commitment could be extended, with a view to adopting a protocol at the third confer-



ence of parties in 1997. The United States is also known to be in favour of starting negotiations on a draft protocol.

This 'worst-case scenario' has caused dismay among environmental groups. Matthew Spencer, atmosphere campaigner for Greenpeace UK, says that if this stance is adopted at Berlin the signatories of the treaty will be no further forward than they were at Rio. Environmental groups and developing countries all interpret the FCCC as committing developed countries stabilizing emissions after the year 2000.

Meanwhile, the United States announced last week its first 'joint implementation' programmes under which it will help to finance a project designed to help reduce the emission of greenhouse gases in a third world country.

Although the principles behind such agreements are opposed by some environmentalists, who claim that they could become an unacceptable substitute for domestic measures, US officials claim that joint implementation projects have "a great potential", and are keen to support such projects in Berlin.

Maggie Verrall

Switzerland seeks greater role in EC research projects

Berne & Munich. Switzerland has opened negotiations with the European Commission (EC) that could lead to Swiss scientists playing a more active role in the EC's fourth Framework programme, which started on 1 January and runs to 1999.

But the talks may not be concluded in time for scientists to participate as fully as many would like, as research is only one of several items in a complex package of issues under negotiation concerning political relations between the European Union (EU) and Switzerland.

Switzerland lost its automatic right to participate fully in the Framework programme when its citizens voted against joining the European Economic Area (EEA) — proposed by the government as a prelude to EU membership — in a referendum in 1992 (see *Nature* 370, 241; 1994).

Last year, the EU decided that Swiss scientists could pay to take part in the fourth Framework programme, but only on a project-by-project basis, and under strict conditions. Unlike scientists in EU states, for example, Swiss scientists cannot become members of the decision-making programme committees. They are also required to team up with partners in two EU member countries to take part in a project, and cannot become project leaders.

"These are big disadvantages for Swiss scientists because they do not have the chance to influence the direction of European research," says Tim Guldemann, a member of the group that advises the Swiss government on science and research policy.

Switzerland is now negotiating to become an associate member of the programme, at a cost of SFr200 million (US\$155 million) a year. This would give Swiss scientists access to the results of all Framework research projects, and the right to initiate and lead research projects with only one EU partner. Switzerland also wants its associate membership to give Swiss scientists the right to participate in the discussions of all programme committees. But this issue remains more controversial.

The EC is insisting that countries that are members of neither the EU nor the EEA cannot have observer or participant status. It has suggested a bilateral committee between Switzerland and the EU at which matters relating only to Switzerland are discussed.

The outcome of the negotiations could be delayed because they are linked to four other issues as a single package. This is bound to slow down the talks, says Jakob Kellenberger, who is leading the negotiations for Switzerland, and who would like each issue to be considered separately.

Oliver Klaffke & Allison Abbott

Sweeping reforms urged for US energy labs

Washington. Comprehensive reform in the way in which the US Department of Energy (DoE)'s 10 national laboratories are managed seems likely to follow the conclusions of a high-level task force, published in Washington last week, that the present system is broken beyond repair.

The task force was chaired by Bob Galvin, chairman of Motorola, and calls for the "corporatization" of the national laboratories. This would detach them from direct supervision by the DoE and allow them to operate like industrial corporations, but under the control of a government-appointed board of trustees.

The task force suggests that legislation that places special restrictions on the laboratories should be repealed, and that funding should be allocated by Congress as separate line items for each of their four main missions, national defence, energy, basic research and environment. It also recommends a gradual reduction in federal funding, claiming that the new structure would make the laboratories between 20 and 50 per cent

more efficient.

The task force's report paints a picture of 10 laboratories filled with competent people and good equipment but overburdened with hundreds of supervisory staff, hundreds of thousands of pages of unnecessary documentation and auditors "descending daily, often by the dozen" to pursue requirements laid down by the DoE or Congress.

As the findings of the report sunk in, some laboratory sources expressed concern that Congress is likely to seize upon the task force's assertion that the laboratories could cost less, and cut their \$6-billion budget without bothering to go through the more exacting process of drafting and passing legislation to reduce the regulatory burden on the laboratories and set up a new management structure responsible for their operation.

But congressional staff on both sides warmly welcomed the Galvin report, predicting that, with strong Republican sup-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Early days: Los Alamos Laboratory has changed dramatically since the Second World War.

port, legislation to reform the governance of the laboratories could be prepared this year and passed in 1996.

"This is a golden opportunity," says John McTague, vice president of technical affairs at Ford, a former acting science adviser to Ronald Reagan and a member of the Secretary of Energy Advisory Board (SEAB), which formally received the report of the task force last Wednesday. "I think a substantial overhaul is called for," says McTague. "But if it isn't done in 18 months it won't be done at all."

But McTague warned SEAB against following the British example — cited in the report — in which the Atomic Energy Authority laboratory at Harwell, Oxfordshire, is being prepared for privatization. Citing his "long experience" of the Harwell laboratory, he says that it has been "unfortunate to see the long-term science base of that facility wither away".

Galvin says that he sees "corporatization" as meaning that the laboratories would be run by "a corpus body separate from government". The government, he suggests, "could own the corporation if it wants to — but it will turn the governance over to the board of trustees". The concept needs refining, concedes the man who ran Motorola from 1959 to 1990, and who still chairs the company's executive board: "If I get some encouragement, I'll do a little study to understand better how it can work."

Hazel O'Leary, the energy secretary, who probably expected a less searing indictment of her department when she appointed the task force a year ago, gave its findings a muted welcome. But she initially rejected the "corporatization" of any of the 10 multipurpose laboratories examined by Galvin. "I see that recommendation as the most extreme in the report," she says. "The single-purpose laboratories are natural for that, [but] we wouldn't submit the 10 big laboratories to such an experiment."

The Galvin findings may be adversely affected by unfamiliarity of many new Republican staff in Congress with the task force, who may therefore ignore its

What the Galvin report said on:

Governance: "The principal organizational recommendation of this Task Force is that the laboratories be as close to corporatized as is imaginable."

The old model "is broken and should be replaced with a bold alternative. Incremental solutions will not likely provide the major improvements that are, at once, achievable and necessary."

The new model: "We do not presume to know what the precise alternative architecture should be One attractive model is a new not-for-profit R&D corporation or corporations ... governed by a Board of Trustees, consisting primarily of distinguished scientists and engineers and experienced senior executives from US industry."

The British example: "Insight should be gained from the experiences of other nations, including the United Kingdom, which recently has maneuvered a disengagement of several of its government laboratories into a semi-privatized status."

Cost savings: "[Governance reform] could be expected to result in an improvement of between 20 and 50 per cent in the effectiveness of the laboratories themselves, on top of significant staff and overhead savings in the Department [of Energy]."

Nuclear weapons laboratories "may not be appropriate candidates for transfer to a non-federal governance structure. Some task force members think they are. Some think they are not."

Basic research: "The task force is con-

cerned about what appears to have been a significant decline in DoE funding for fundamental research over the past three years, with the prospect of still deeper cuts to follow. The department should sustain and strengthen its support of fundamental science."

Industrial competitiveness: "The industrial competitiveness activities at the national laboratories are unfocused and lack a firm policy foundation. Industrial competitiveness, broadly defined, has no place as a stand-alone mission of the laboratories The idea that the laboratories are, or could become, cornucopias of relevant technology for a broad range of industries is a myth."

Environmental management: "the bulk of environmental challenges ... are becoming calcified, and the vast flow of funds into the program acts like an anesthetic, numbing the Department, State regulatory agencies and affected stakeholders, hindering and delaying beneficial change."

Lawrence Livermore National Laboratory (LLNL): "LLNL should retain enough nuclear weapons design competence and technology base to continue its activities in non-proliferation, counter-proliferation, intelligence support, and verification, and to provide independent review [of work at Los Alamos]. LLNL would transfer, as cost-efficiency allows, over the next five years its activities in nuclear materials development and production to the other design laboratory [Los Alamos]."

►fiscally and ideologically attractive advice. Supporters of the findings hope that laboratory directors — who dislike existing DoE oversight arrangements, and have good congressional connections — will help win support for radical reforms.

But Nick Samios, director of the Brookhaven National Laboratory, New York, says that reform could work within the existing, government-owned, contractor-operated (GoCo) system. "I know that a real GoCo will work," he says, claiming that 'corporatization' could damage the laboratories. "I'm not against it," he says. "I just want to see what it is."

Bruce Tarter, director of the Lawrence Livermore National Laboratory in California, which Galvin said should lose its nuclear weapons design capability, issued a statement predicting that Galvin's recommendations "will turn out differently" when evaluated in the context of the entire nuclear weapons programme.

Others suggest even more radical steps. An "Alternative Galvin Report", published by the Military Production Network (MPN), an environmentalist coalition, urged Congress to save \$4 billion over four years by simply scrapping nuclear weapons research altogether.

"Continued weapons research, development and testing is contrary to stated national goals of 'reducing the global nuclear danger,'" MPN says. "It's time to end the Manhattan Project." **Colin Macilwain**

Gene therapy trial passed in Japan

Tokyo. Japan's first clinical application of gene therapy was approved earlier this week by a committee of the Ministry of Health and Welfare (MHW).

A research team led by Yukio Sakiyama of Hokkaido University School of Medicine was given permission by the ministry's central assessment committee for gene therapy to carry out clinical research in gene therapy in order to treat a four-year-old boy suffering from adenosine deaminase (ADA) deficiency, an enzyme deficiency that weakens the immune system.

The permission has been given on condition that Sakiyama makes some revisions to the forms used to obtain informed consent from the patient, and that he explains the treatment once more to the boy's parents.

The treatment he plans to use is based on a technique developed five years ago at the US National Institutes of Health (NIH). It involves introducing the ADA gene to the child's T-lymphocytes using a bacterial gene for neomycin resistance.

Sakiyama had initially proposed a slight modification of the technique. But during the approval process this idea was dropped for safety reasons, and he will now use an identical technique to that

UK shifts research support into 'strategic' initiatives

London. The British government announced last week that it is to increase funding for 'strategic' science initiatives and for collaboration with industry by £67 million in the financial year starting on 1 April, equivalent to around five per cent of the total budget of £1,281.7 million.

Overall, however, the science budget will increase by 3.3 per cent next year, slightly higher than the anticipated rate of inflation. The emphasis being given to strategic initiatives will inevitably mean less money for areas of basic science, and is already causing difficulties in areas such as astronomy (see page 459).

In announcing the allocation of the science budget for 1995-96, David Hunt, minister for science, said that the support for strategic initiatives is intended to promote policies highlighted in the government's science and technology white paper of May 1993, *Realising Our Potential*.

The money will be used to extend existing schemes to improve interaction with industry — such as the Realising Our Potential Awards (ROPA) scheme — to increase funding for 'underpinning' strategic science in physics, mathematics, medicine and biol-

ogy, and to enhance 'people-related programmes', including the promotion of the public understanding of science.

Sir John Cadogan, director general of the research councils, has 'ring-fenced' £8 million of the science budget to insulate the Particle Physics and Astronomy Research Council (PPARC) from fluctuations in its international subscriptions due to changes in exchange rates. PPARC had complained that increases in subscriptions to the European Laboratory for Particle Physics (CERN) and the European Space Agency had left it unable to fund instrumentation and scientists to exploit the contributions. Cadogan said the chief executives of the other councils agreed to the ring-fencing, but were watching the money "like hawks".

The ROPA scheme, which funds researchers who are already working in collaboration with industry, will be one of the main beneficiaries of the new budget, with additional funding of £14.95 million that will allow it to extend to all six research councils. Cadogan says the outcome will be "blue-skies curiosity-driven research" that will also benefit strategic research.

Another winner in the budget will be the LINK scheme, soon to be relaunched with £3 million earmarked from the research councils' budgets. The Medical Research Council (MRC)'s genome research programme gets a £3.5-million fillip.

Chemistry is to retain enhanced support through the Biotechnology and Biological Sciences Research Council (BBSRC), the Engineering and Physical Sciences Research Council (EPSRC) and the Natural Environmental Research Council (NERC), at a total of £7.6 million.

Physics, mathematics and medicine benefit from £4.4 million allocated for strategic initiatives, and the BBSRC will receive an extra £2 million to support research that underpins emerging technologies in bioprocessing and developing 'wealth-creating products' from plants, such as vaccines and biopharmaceuticals.

The Royal Society will also benefit from an increase in the number of its research fellowships from 200 to 255. Extra funding will also allow it to introduce new fellowships to promote the role of women in science, named after the late Dorothy Hodgkin.

The lack of overall growth in funding has already brought criticism from groups such as the Institute of Professional, Managers and Specialists, the labour union which represents many research council staff. But Hunt described the budget package as "a vote of confidence in the country's engineers, mathematicians, natural and social scientists". **Maggie Verrall**

already successfully used by Michael Blaese of NIH.

Because the retroviral vector to be used in the treatment is not available in Japan, Sakiyama has to obtain approval from the US Food and Drug Administration for the export of the vector by Genetic Therapeutic Inc. of Maryland. As a result, he is not expected to be able to begin treatment until about April.

The approval this week was the last step in a complex process overseen by the MHW committee and another separate committee in the Ministry of Education, Science and Culture. That committee, which was concerned with scientific and safety issues, gave its approval last week.

The MHW committee meeting, which unusually for Japan was open to the press, concentrated on the issue of informed consent, for which Sakiyama is using detailed forms based on a translation of those being used by Blaese in the United States.

There is still little awareness among Japan's medical profession of the Western concept of informed consent, and one additional benefit of the application of gene therapy may be a more rigorous approach to this issue in general medical practice. **David Swinbanks**

University challenges plagiarism judgement

Tokyo. The University of Hong Kong, rejecting the judgement of Hong Kong's High Court and Court of Appeal, has cleared a faculty member of charges of plagiarism and infringing copyright in a questionnaire used for an epidemiological study of lung cancer. But the impartiality of the university's investigation has been questioned by the researcher who brought the initial charges.

In a precedent-setting court case, the appeal court confirmed in August 1993 a high-court verdict reached the previous year that Lam Tai Hing, a reader in the university's department of community medicine, made unauthorized use of a questionnaire devised by a lecturer in the department (see *Nature* 366, 715; 1993). But a subsequent internal inquiry by the university cleared Lam of the charges.

The case started in 1986 when Linda Koo, a lecturer in the department, accused Lam of copying parts of a questionnaire that she had devised with John Ho Hung Chui, honorary professor of radiation oncology at the university, to study the incidence of lung cancer in non-smoking women.

Her claims set an unusual precedent, as several leading epidemiologists argue that such questionnaires are usually freely circulated among researchers and are not considered subject to copyright. But Koo alleges that she was being harassed by Lam and the then head of the department, J. W. I. Kleevens, and that this was a driving force behind her decision to go to court.

The university's inquiry was carried out by a committee of two pro-vice-chancellors and four members of the senate. The committee met in private 14 times between October 1993 and November 1994 and heard evidence in Lam's defence from Lam, his lawyer and several witnesses.

It also examined some of the extensive court records and received written opinions on the similarity of Koo and Lam's questionnaires from 26 eminent epidemiologists from Britain, the United States, Singapore and Australia. Neither Koo nor Ho, their lawyers, nor witnesses for the plaintiffs, participated in the hearings.

In a report released on 20 January, the committee, saying that it required more substantial proof of guilt than the high court, dismissed all charges against Lam. It claims that new evidence shows that an assistant of Koo's may have given the questionnaire to Lam's assistant, and concludes there is no evidence that Lam stole it, as Koo alleges.

In contrast to the court ruling, the committee says that Lam's questionnaire was not copied from Koo and Ho's. It also quotes a "clear majority view" from epidemiologists and other scientists that in the 1980s "there was no copyright or confidentiality in questionnaires, and the practice was free exchange and use of such material".

But Koo challenges the impartiality of these experts. She points out that their comments were solicited by the head of Lam's department, A. J. Hedley, and that, in the words of the committee's report, Hedley's letter to them was "highly tendentious and was (in lawyer's language) 'leading'".

For example, Hedley's letter said that Koo's attitude was affected by her failure to gain promotion in 1984, questioned the basis of the judge's decision, mentioned a widespread sense of grave miscarriage of justice and also spoke strongly of the need for free circulation of questionnaires. Nevertheless, the committee said it placed great weight on the views of these 26 experts as they are "senior and distinguished scientists, who can be expected to have an independent mind".

Richard Peto, professor of medical statistics and epidemiology at the University of Oxford in Britain, who acted as expert defence witness in the high court trial and also presented evidence to the internal inquiry, says "justice has been done at last". In

contrast, Koo says the inquiry was a "kangaroo court" and "whitewash job" to clear Lam's name and the reputation of the university. She says the committee report shows "total disrespect for the court system".

The committee criticizes Koo and Ho for not agreeing to give oral evidence and denying the committee access to relevant material. But Koo wrote to the committee's secretary suggesting dates on which she and Ho could appear before the committee, and she offered to bring all 75 box files they had on the case to the hearing. The committee replied that it "did not initially foresee the need" for them to bring the box files. It also said that, while Lam would be represented by his lawyer, who could ask questions of them and other witnesses, the lawyer for Koo and Ho could attend only as an "observer". On advice from their lawyer, the two declined to participate.

The university says it considers the whole matter closed. But Koo says that she is now seeking advice from her lawyer on the possibilities of legal action against the university. **David Swinbanks**

Darwin's islands under threat

London. Scientists are being asked to write letters of protest to the President of Ecuador about threats from sea-cucumber fishermen to the wildlife of the Galapagos Archipelago, the islands that gave Darwin the inspiration for his theory of natural evolution.

According to conservationists, the overfishing of the sea cucumber, *Isostichopus fuscus*, an important link near the base of the food chain, and an influx of people to the islands to exploit the lucrative trade, is threatening the wildlife of the islands. This in turn could destroy the tourist industry on which many of the 12,000 islanders depend.

Sea cucumbers are marine invertebrates that lie on the bottom of the sea feeding on detritus and plankton. The eight-inch-long rusty brown creatures are sought after in China and South-East Asia for their supposed medicinal properties.

Two months after the opening of a fisheries in the islands last October, fishermen were reported to have taken more than seven million sea cucumbers, despite a quota of 500,000 in a three-month season set by the Ecuadorian government. Snails, sea horses, sea urchins and black coral are also reported to have been taken.

The Ecuadorian authorities closed the fisheries in December, sparking violent protests last month from fishermen and dealers. According to Macarena Green, a scientist on the islands, protesters blockaded the

Charles Darwin Research Station and the headquarters of the national parks service. They took hostages and threatened to kill the giant tortoises at the research station.

In an attempt to pacify the protesters, a government official granted them an extra 30 days of fishing. But this was overturned almost immediately and replaced with a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

At risk: one of the Galapagos' giant tortoises.

nine-month ban after a meeting between the minister of fisheries, fishermen and conservationists. The ban reflected the efforts of scientists working with local environmental groups, tour operators, travel agents and the association of ecotourism to persuade the minister that the activities of the fishermen pose a threat to the islands.

In a message that has been broadcast on the Internet, Green urges scientists to bring international pressure on the Ecuadorian government to close the fisheries permanently. **Maggie Verrall**

Europe dilutes draft bioethics convention

Paris. The parliamentary assembly of the 32-member Council of Europe last week approved the draft European Convention on Bioethics. But the agreed text, which had been considerably watered down in comparison to an earlier version after opposition from Germany, faces a long haul before being approved by the Committee of Ministers and ratified by member states.

Indeed, the difficulty in approving the text — which is being billed as the first international treaty specifically to address the potential risks of progress in medicine and biology to the rights and freedoms of individuals — has highlighted the impossibility of achieving international consensus on bioethical issues.

Germany, for example, refused to withdraw its opposition to an article of the convention, which would have permitted research on 'non-viable' embryos that had not developed beyond 14 days. The new text leaves the question open by deleting any reference to embryo research, apart from banning embryo production for research purposes. "Germany didn't want to hear the term embryo research," say council officials.

Also at Germany's insistence, the new text would ban all genetic research on 'legally incapacitated persons', which would include both handicapped persons and children. But such research, which is essential to study many genetic diseases, is already permitted in countries such as France under strictly regulated conditions.

Changes to the initial draft of the convention were sufficient for it to obtain the approval of a majority of the council's parliamentary assembly meeting in Strasbourg last week. But as its provisions on legally incapacitated persons conflict with existing

national legislation, these will have to be reworded before the convention is voted on by the Committee of Ministers. France, for example, could not legally adhere to the convention as it now stands.

Council officials say that the textual changes needed to resolve this dilemma will probably require the parliamentary assembly to vote on the convention again later this year. If so, the Committee of Ministers would not be able to vote on the text before around this time next year.

The outcome of such a vote is far from certain. Italy, for example, where the Roman Catholic Church has strong views on

bioethics, last week voted against the whole text. Adoption of the convention by the ministers only requires a majority vote. But moves to open it for ratification by the member states can be vetoed by individual states.

Council officials claim that neither Germany nor Italy are likely to veto the convention. They argue that both countries will probably accept that any convention is better than nothing, and at worst opt out by refusing to ratify the text themselves. Formal adoption of the convention requires only that it be ratified by five member states, after which it becomes legally binding in all those states who ratify it. **Declan Butler**

Germany eyes new ozone controls

Munich. Germany's federal environment minister, Angela Merkel, has stepped up efforts to introduce nationwide laws to reduce environmental ozone concentrations by banning cars without catalytic converters on sunny summer days.

Her move follows the publication of Europe's first study of the effects of local traffic restrictions in a small geographical area, which indicated that ozone levels can be reduced only if controls are more widely applied.

The European Commission wants to go even further. It has proposed Europe-wide measures to be applied to both traffic and industrial emissions. A directive reducing ozone emissions from industrial and other sources was adopted by the European Union (EU) last December, and the commission is now working on further measures.

Ozone levels have been rising in all industrialized countries over the past

decade, and often exceed the warning levels contained in the EU directive of 360 micrograms per cubic averaged over one hour, in all of Germany's 16 *Länder*, one of the highest reported incidences of excessive ozone levels in Europe.

The recent test was carried out by the environment ministry in Baden-Württemberg. Cars without catalytic converters were kept off the roads in two neighbouring cities for four sunny days last summer. The tests showed that levels of NO fell by 30 per cent and volatile organic compounds (VOCs) by 15 per cent.

Both chemicals are involved in the formation of ozone, and catalytic converters reduce the release of NO and VOCs by 95 and 99 per cent respectively. But local ozone levels did not fall as neither chemicals were present at saturating concentrations.

Harald Schäfer, minister of the environment for Baden-Württemberg, argues that the results show that, to reduce ozone levels, traffic restrictions need to be implemented over a much larger area.

Merkel agrees, and wants to introduce federal laws giving individual *Länder* the freedom to restrict traffic when there is a risk of summer smog. Last November, Merkel and the regional ministers of the environment unanimously supported a proposal to change Germany's Air Quality Act to allow this.

If the proposal is accepted, cars without catalytic converters and diesel cars with high pollutant emission levels will be banned from travel on high-risk days. But it has not yet been decided if the same restriction will be imposed on lorries, or whether industrial emissions will also be subject to control.

But the federal ministry of the environment is not likely to demand that speed limits be imposed. There is no evidence that such measures significantly affect ozone levels, and the issue is politically sensitive, as speed limits would be hotly opposed by the transport ministry. **Toni Feder**

MRC concordat 'under threat', say Lords

London. Members of Britain's House of Lords voiced concern last week that future activities of the Medical Research Council (MRC) could come under threat if a 'concordat' between the MRC and the National Health Service (NHS) is substantially changed when it comes up for renewal later this year.

The agreement, signed in 1992, promises that health departments will cover the infrastructure and overhead costs of research projects supported by the MRC. But a task force set up by the Department of Health to look at support for research and development in the NHS said in its report last year that the agreement was "too open-ended" (see *Nature* 371, 275; 1994).

Giving evidence to a subcommittee of the Lords Select Committee on Science and Technology, Anthony Culyer, the chairman of the task force and professor

of medical economics at the University of York, said the current arrangement amounted to a "blank cheque" for the MRC. "It didn't seem right to us that the written concordat says that the NHS would support whatever research the MRC chooses."

The task force did not recommend changes, but, according to Lord Perry of Walton, a member of the committee and a former member of the MRC, this is implied by the task force's conclusions.

But Lord Walton of Detchant, chairman of the subcommittee — and also a former member of the MRC — said it was "perfectly reasonable" for the MRC to determine research priorities. He added that there was a "sufficient mechanism" in place at present for the health department to make its wishes known to the MRC through representatives on the research council's committees. **M. V.**

Italy pursues cuts in manned space efforts

Rome. The Italian Space Agency (ASI) is planning to seek reductions in its various commitments to the space programmes of the European Space Agency (ESA) in an effort to draw its books into balance its books.

Italy's move will add to ESA's troubles in getting agreement from its member states on its proposed ECU3.9-billion (US\$4.8-billion) budget for the next eight-year (1995–2003) manned space flight programme, which is already under attack from both Germany and France.

European manned space efforts have been losing political support in the 1990s as recession has hit ESA's member states. Italy in particular has faced severe difficulties, while ASI itself has been troubled by administrative and financial problems since it was set up in 1988.

In the early 1980s, Italy decided to become a major force in European space. But by the 1990s it was no longer prepared to back up this political will with the necessary financial commitment. ASI's budget, for example, has been frozen for the last three years, and by last summer the agency had already run up an overall deficit of IL500 billion (US\$310 million).

Most of the ASI budget is used to support Italy's participation in ESA programmes. Even so, in 1993, it was able to cover only two-thirds of its ESA bills, and early last year, ESA agreed to take out a IL189-billion loan on ASI's behalf, to be paid back by the Italian government before the end of 1996.

Meanwhile, research scientists in Italy were protesting that the agency was diverting funds that they claimed were intended for the national basic research programme to support its ESA commitments (see *Nature* 356, 647; 199).

Two years ago, Umberto Colombo, research minister in the interim government, installed a commissioner to try to sort out its problems. By last summer, the agency was back on its administrative feet, with a new president and a new board. In particular, it was once more able to distribute grant money to research scientists, for example.

The new director general, Mario Calamia, has since been renegotiating international commitments, particularly those with ESA. The heaviest commitments, which include data relay satellite programmes and early developmental parts of manned space flight programmes, run until 1996. Calamia says that Italy will not backtrack from these, but will instead try to delay payments, possibly by securing another bank loan through ESA.

Most important, he says, is to ensure that post-1997 commitments are reduced. "We want to extend the timeframe of some commitments and we will certainly cancel others", he says. He will not be drawn on where the axe might fall, but it will almost certainly involve its obligations to the ESA

manned space flight programme, to which Italy committed itself heavily at the ESA ministerial meeting in Grenada in 1992.

The original manned space programme, promoted heavily by Italy in collaboration with Germany in the mid-1980s, was intended to guarantee Europe autonomous access to space, as well as the opportunity to participate significantly in the US international space programme.

It included a small space laboratory known as the Man-Tended Free Flyer (MTFF), a space plane called Hermes and an attached pressurized module (APM), part of the so-called Columbus programme, which would attach either onto the MTFF space station or the planned international space station.



Calamia: seeking to balance the books.

But, given Europe's changed economic circumstances, the MTFF and the Hermes programme were formally abandoned at the Grenada meeting. The Columbus programme was given the green light.

But in the last two years the APM has been scaled down to a much smaller unit called the Columbus Orbital Facility (COF) which will dock with the international space station, and which should be around half the price.

Last month, ESA put forward a proposed ECU3.9-billion manned space flight programme, ECU1 billion of which would be earmarked for COF. The proposed budget has to be approved at the next ministerial meeting in November this year. As a result, Italy, which enthusiastically agreed to fund 31 per cent of the Columbus programme in Grenada, could be asked to pay over ECU300 million, between 1995 and 2003.

Italy might try to withdraw from the programme altogether. But that would put its partners under severe pressure. Italy is more likely to seek a smaller share of the programme, and ask that it be extended over a longer period to reduce annual cost.

The second option in particular could win support from Germany and France,

ESA's largest contributors. They have already responded to ESA's proposal by saying that they will not support a figure higher than ECU2 billion for the eight-year programme.

Calamia says he does not intend to join Britain's fight to reduce the costs of ESA's mandatory space science programmes (see *Nature* 372, 712; 1994). "The money involved [IL97 billion in 1994] is very small compared with our commitments on optional programmes [IL419 billion]", he says.

In addition to trying to reduce ESA payouts, Calamia is exploring possible ways of reducing ASI's space commitments under bilateral agreements with the United States. The heaviest of these is its agreement to provide three Mini-Pressurized Logistic Modules (MPLMs) to carry cargo up to the space station, at a cost of US\$500 million. The launch of the first MPLM is due in December 1997.

But Calamia says that ASI will not curb its spending on the controversial X-ray satellite SAX. This is a national programme that has suffered excessive price increases and extended delays, which, many scientists argue, have reduced its scientific value so much that it is no longer worthwhile (see *Nature* 366, 101; 1993).

ASI is not being helped in solving its financial problems by continual upheaval in its administration. Former research minister Stefano Podestà, who succeeded Colombo last summer, twice tried unsuccessfully to force through emergency legislation dissolving the agency's administration once more and installing an *amministratore unico*, a 'single administrator', with full control of all ASI business (see *Nature* 372, 719; 1994).

Podestà finally succeeded in pushing through a decree during the turmoil of the last days of his government last month. But while the decree immediately removed the administration, no *amministratore unico* was appointed, and the agency was left head-less for a month.

Last week parliament rescinded the decree, and the former administrative bodies have been reestablished. Parliament will now consider a bill with the same aims as the decree — namely to install a single administrator with broad powers — but which, unlike Podestà's bill, will be fully debated.

Alison Abbott

Chile court backs locals in telescope row

Munich. Chile's supreme court last week ruled in favour of the Latorre family which claims to own the land on which the European Southern Observatory (ESO) is currently building its DM463-million (US\$300-million) Very Large Telescope (VLT).

As a result, work on the site, whose ownership has been the subject of an intense and protracted legal battle (see

Nature 368, 676; and 370, 494; 1994), is being required to stop until the claim is resolved.

But the ESO has already invested heavily in construction work at the site on the top of Mount Paranal in northern Chile, and is continuing with the building work as it has not received any formal instructions from the government to stop. A. A.

Revolutionary birthdays

SIR — A number of reports have linked season of birth with psychological aspects as diverse as mental illness¹ and occupational choice². Recent correspondence³⁻⁵ concerning birth data effects has prompted this report of a relationship between season of birth and stance taken in the scientific revolutions associated with the theories of relativity and evolution.

To ascertain whether there were significant differences in birth dates between proponents and antagonists of relativity theory, the biographies of physicists who took and maintained a committed position from an early stage were scrutinized. The groups were well matched for eminence: 5 of the 10 relativists and 4 of the 9 non-relativists were awarded the Nobel prize.

Investigation of birth dates revealed that 8 of the 10 relativists were born during December, January and March, whereas 6 of the 9 anti-relativists arrived during June and July. Contrasting the months of October to April and May to September, all 10 relativists were born in the winter months as against only 2 of the 9 anti-relativists ($P < 0.001$, Fisher test).

A similar investigation was undertaken with those eminent biologists who advanced or opposed evolutionary theory before publication of *The Origin of Species* in 1859. Again the differences were marked, with October to April accounting for the births of 11 of the 12 evolutionists but only 5 of the 16 anti-evolutionists ($P < 0.002$, Fisher test). Aggregating both scientific debates, December to April houses 82 per cent of the combined proponents' birth dates but only 24 per cent of those of the antagonists; in contrast, May to July accounts for none of the proponents' but 60 per cent of the antagonists' births ($\chi^2 = 18.0$, $P < 0.001$).

One hypothesis suggests itself. Different dates of birth will result in the seasons being experienced at different stages of early development; and the concept of critical or sensitive periods of development suggests that the timing of such experience may have far-reaching effects. Before birth, the developing embryo or fetus is influenced indirectly through maternal behaviour and biology; and maternal activity, diet and health may all change with the seasons. Light-induced hormone fluctuations may be a particularly potent influence on early development.

The social environment of the infant may also be heavily influenced by the seasons. To return to the scientific revolutionaries and non-revolutionaries, all were born in Europe or North America, on average a little north of 50° latitude. Without electric light and central heating, the seasons would have affected mother-infant interaction. Initially the winter-born revolutionary would have been res-

trictively swaddled but with the coming of summer he would have experienced more freedom to explore on his own initiative. The summer-born non-revolutionary, on the other hand, would have enjoyed more freedom at first, when less able to use it, but would have been constrained the following winter when ready to extend his explorations. Could such early experiences have shaped attitudes toward established paradigms?

Michael Holmes

*Psychology Section,
Queen Margaret College,
Edinburgh EH12 8TS, UK*

1. Hare, E. *et al.* *Br. J. Psychiat.* **124**, 81-86 (1974).
2. Cooper, J. & Smithers, A. *New Society* **24**, 6-8 (1973).
3. Dudink, A. *Nature* **368**, 592 (1994).
4. Edwards, S. *Nature* **370**, 186 (1994).
5. Baxter-Jones, A. & Helms, P. *Nature* **370**, 186 (1994).

ESF membership expanded

SIR — Your report (*Nature* **372**, 395; 1994) stated that the statute of the European Science Foundation was amended "to remove the possibility that countries such as those of central and eastern Europe might acquire associate status". In fact, the assembly in November decided to encourage qualified science funding organizations in central and eastern Europe to apply to proceed directly to full membership of the foundation.

Following this decision, the foundation has begun to develop contacts with central and eastern Europe with the expectation of extending our membership from that area.

Peter Colyer

*European Science Foundation,
1 quai Lezay-Marnésia,
67080 Strasbourg Cedex,
France*

Coelacanth dated

SIR — I am as excited as anybody by the discovery of the Wollemi pine (*Nature* **372**, 712 & 719; 1994) — I love "primitive" gymnosperms, and I think it is indeed comparable to the discovery of the first living coelacanths. But *Latimeria chalumnae* was discovered in 1938 surely, not 1948 — some of us still remember its announcement as among the few pieces of wholly pleasant scientific news which came out as the war-clouds were gathering over Europe.

Oliver Sacks

*Department of Neurology,
Albert Einstein College of Medicine,
Bronx, New York 10461, USA*

Japanese research

SIR — A recent News article (*Nature* **371**, 734; 1994) wrongly stated that in Japan, "university researchers are prohibited by law from participating in research projects in industrial laboratories". Researchers from national and international universities conduct research in our laboratory, and this type of collaboration is not only legal, but encouraged. Professors in Japan, as public-sector employees, are not allowed to work as managers in companies and must seek permission from their universities before earning outside income. Although the tone of the News article was critical of the Japanese system, a leading article in the same issue (**371**, 725-726; 1994) suggested that to help US and British university researchers working in industry avoid conflicts of interest, "a cap on consultancy and other earnings would be only seemly". If Japanese academics truly are not "pushing for regulations that prevent them from working in company laboratories to be erased", then this seems laudable, not cause for censure.

Glen Brown

Shiono Satoru

*Central Research Laboratory,
Mitsubishi Electric Corporation,
Amagasaki, Hyogo 661, Japan*

Moral premises

SIR — Tom Gehrels' forthright review of the apologia for Werner von Braun (*Nature* **372**, 511; 1994) raises the important question of the moral status of curiosity-driven science and technology. Awe at the heroism of those who attempted to sabotage Braun's rocketry is inevitably reinforced by distaste for the man himself and those who did not scruple to provide him with means to carry on with his work during and after the Second World War. A psychopath remains a psychopath however clever; his talents just make him more dangerous and reprehensible. Rocketry has, no doubt, facilitated astronomical research into the nature of the Universe, but its main fall-out has been to make possible the dissemination of fourth-rate and often corrupting entertainment.

What the whole story makes clear is the urgent need for scientists to examine the moral premises underlying their choice of research — not so much the possibility that their work may be put to evil as well as good use, but their own motivation and its place in the activities of a humanist university or an industry supposedly catering for legitimate human needs. It is only too easy to make a Faustian bargain with the devil.

John A. Davls

*1 Cambridge Road,
Great Shelford, Cambridge CB2 5JE, UK*

Classical and quantum physics mix

A neat way of talking of quantum and classical physics in the same language will be of interest in itself even if it does not avoid the need for a quantum theory of gravity; but that would be a huge extra.

MIXING together classical mechanics and quantum mechanics is a long-standing source of embarrassment, usually arising quite early in the textbooks of quantum mechanics. There is, for example, the simplest of all the problems in quantum mechanics, that of a particle in a box. The heuristic value of this example is that it stands proxy for an electron or some other charged particle in a fixed electrostatic potential, as in the interior of a metal for example.

Beginning students are not, however, invited to wonder how the electrostatic box itself is constructed. If they were (and if the dimensions of the box were of the same order of magnitude as the wavelength of the electron), they would quickly conclude that the electrostatic field that constitutes the box can have been made only from microscopic objects, say the ions of a crystal lattice, which are themselves, in principle, also subject to the laws of quantum mechanics. So why treat the electron in a box as a quantum object, yet consider the electrostatic potential that confines it as a classical given?

Similar problems arise in more taxing fashion in textbook discussions of the problems of quantum measurement. The Heisenberg *gedanken* microscope used to demonstrate the inescapable uncertainty in a measurement of the position of an electron is a case in point. The idea is to estimate the position of the electron from the scattering of photons. The precision of the measurements is inevitably a function of the frequency of the photons, but the greater the frequency, the greater the momentum they will transfer to the electron, which offsets the potential benefit of working with higher frequencies.

The argument demonstrates what the authors of the textbooks want to show, that Heisenberg's Uncertainty Principle is unavoidable, but hardly anybody stops to ask whether the microscope itself is a quantum object. If, by chance, they do, the answer will be that it is a macroscopic object, the uncertainty in the position of its optical elements (calculable from the same uncertainty principle) will be shown to be much less than the uncertainty in the position of the electron, and that will be that.

Thus the world has become used to the idea that quantum physics and classical physics are a little like oil and water. They do not mix. Where some interaction is of the essence, as in the behaviour of a laser, the quantum elements (excited atoms and the

photons whose emission is stimulated) are dealt with as such, while the structural elements of the laser (which define the shape of the resonating cavity) appear simply as parameters in the equations. This fuss does not imply that the standard treatments of mixed problems are in some sense wrong, but that they are at least inelegant and may leave people with an awkward sense of insecurity.

That seems to be part of the reason why Arlen Anderson, from Imperial College London (but also affiliated to the Newton Institute at Cambridge) has embarked on a scheme for including both quantum and classical variables in a common theory (*Phys. Rev. Lett.* **74**, 621–625; 1995). As he puts it, he seeks to “assuage mathematical doubts about how classical and quantum variables may coexist in a single theory”. But reading between the lines, the long-term objective is to find a way of handling the field theories of particle physics self-consistently against the background of a classical theory of gravity, that of Einstein's general relativity.

Anderson acknowledges earlier work in the same direction by I. V. Aleksandrov and by W. Boucher and J. Traschen, saying that the distinctive contribution is to demonstrate the ‘back-reaction’ of the quantum system on a classical system with which it is coupled. This effectively takes the form of an extra source of noise in the evolution of the classical system whose effect is to induce a correlation between the values of classical observables (the quantum word for ‘measurable quantity’ or ‘variable’) and those attained by the coupled quantum system.

The algebraic technology of putting together quantum and classical mechanics turns out to be simple enough, perhaps deceptively so. The starting point is the definition of a conjugate pair of variables (say $\mathbf{q}$ and $\mathbf{p}$, perhaps the position of a particle and its momentum in that direction), to which the Uncertainty Principle applies in quantum mechanics, which is that $[\mathbf{q}, \mathbf{p}] = i$ (where $[\mathbf{q}, \mathbf{p}] \equiv \mathbf{qp} - \mathbf{pq}$, $i \equiv \sqrt{-1}$ and the units are such that $\hbar/2\pi = 1$).

The classical analogue of this relationship is familiar from Hamilton's nineteenth-century formulation of classical mechanics, which makes it possible to test whether, given one pair of conjugate variables, say x and k , functional combinations of them, say $f(x, k)$ and $g(x, k)$, are also conjugate variables. The test is that the so-called Poisson bracket, called $\{f, g\}$, should be equal to 1.

The Poisson bracket is defined as $(\partial_x \partial_k g - \partial_k \partial_x f)$, where ∂_x and ∂_k denote partial differentiation with respect to x and k respectively. (Plainly the test is trivially satisfied for the original variables x and k .)

What Anderson does is simply to put the two conditions together, arriving at a test for pairs of conjugate variables (A and B) with the form $[A, B] + i\{A, B\} = i$. At the most trivial level, this equation is plainly valid; if A and B are a conjugate pair of quantum variables, only the first term applies (and equals i by definition), but if they are a classical pair, only the second is relevant and is identically true.

More generally, he defines a new bracket expression for any two variables as $[A, B]^*$, say, by the expression above, $[A, B] + i\{A, B\}$; he departs from others who have followed this route in the rules for evaluating $[A, B]^*$ for mixtures of classical and quantum variables. Following the standard procedures of quantum mechanics, there then tumbles out an expression for the time rate of change of an arbitrary function A of quantum and classical variables (and in which the time is not explicit) in the familiar form of $-i[A, H]^*$, where H is the Hamiltonian of the system.

The physics of this formalism is not as obscure as it may seem (but Anderson's paper illustrates the point with two workable — and neatly worked — examples). The deficiency of representing the variables of the classical parts of a mixed system as parameters in quantum equations of motion is that they are thereby given the status of givens, and cannot be changed. But even if the classical parts are so far away from the conditions required of quantum systems that they will not, for example, be quantized in the sense that only some states are accessible to them, there is every reason why their continual interaction with a quantum part will affect the probability distribution of the states in which they may be found.

Whether these developments will solve the problem of the gravitational field is another matter. After two decades of serious attempts to produce a quantum theory of gravitation, it is forgivable that people should now be wondering whether it is really necessary to try. The clamant need for quantum gravity arises only in exceptional circumstances, at the edges of black holes and at similar exceptional locations. But to follow even Anderson's new approach, people will have to learn to do quantum mechanics in curved space-time. That will not be child's play.

John Maddox

How special is Jupiter?

George W. Wetherill

BUILDING Jupiter has long been a problem to theorists. Observations presented on page 494 of this issue¹ confirm that the problem is indeed a serious one.

The idea that stars and planetary systems including our Solar System form simultaneously from a disk-shaped nebula of gas and dust is an old one, dating back to the days of Kant and Laplace. In recent decades this hypothesis has finally received considerable observational and theoretical support. Infrared, millimetre and optical wavelength observations of very young stars demonstrate the presence of infalling circumstellar disks of dust and gas associated with these stars. Theoretical calculations have supported plausible mechanisms whereby, in a symbiotic manner, the excess angular momentum associated with a 'core' of a parental interstellar molecular cloud may be partitioned between the newly formed central star and a residual disk of gas and dust. The star thus obtains the mass it needs to function as a hydrogen-burning thermonuclear furnace, while losing the excess angular momentum that would prevent it from collapsing into a compact body. Much of this angular momentum, associated with only a small part of the mass, can be transferred to a residual centrifugally supported circumstellar disk from which planets may form.

At present, our planetary system contains no more than a few per cent of the gas, primarily hydrogen, that should have accompanied a preplanetary disk of cosmic chemical composition. How and when was this excess gas lost? Our well behaved middle-aged Sun is not capable of 'blowing away' such large quantities of gas. It seems most likely that the necessary loss of gas from the plane of the disk would occur during the less stable first few million years of a star's life.

Zuckerman *et al.*¹ report radiofrequency observational data supporting this picture for 20 nearby young stars, all estimated to be between 1 and 10 million years of age, using methods of varying reliability. They determined the gas density from the intensity of carbon monoxide molecular transition lines, direct determination of H₂ being impossible because of the absence of suitable H₂ radiofrequency transitions. Concentrations of H₂ were then obtained by assuming the standard interstellar ratio (10⁴) between the number of H₂ and CO molecules. Subject to many caveats, the authors conclude that within a few million years after the formation of a star, the mass of H₂ gas remaining is typically very small, much less than the mass of Jupiter. An exception to this has been found for the still

young star GM Aurigae², which is roughly 2 million years old.

This early loss of gas could make problems for the formation of Jupiter and Saturn. Theoretical models for the formation of Jupiter usually first require the growth of a core roughly 10 times the mass of the Earth, upon which the gas necessary for Jupiter to grow to its full 317 Earth masses will accrete from the disk in a few million years. The requirements for the formation of Saturn are similar. Obviously this entire process must take place before the gas has been lost from the disk. It has proved awkward to develop a quantitative model that permits a sufficiently large Jupiter core to grow during the short time that gas is present, and at the same time is consistent with the absence of large planets in the asteroid belt, although this may be possible for seemingly *ad hoc* distributions of mass of gas and dust in the disk. An 'easy' way out of this dilemma could be to hypothesize that the special conditions necessary to form Jupiter and Saturn represented an unusual stochastic circumstance, and that planetary systems rarely include gas giant planets, even if their terrestrial planet regions are populated by bodies generally similar to those of our Solar System.

There is some observational support for this. Conventional models for the formation of the 'Oort cloud' of comets, roughly 10⁴–10⁵ astronomical units (AU) distant, predict that this process is inefficient and that many more comets were ejected from the Solar System than were preserved in the Oort cloud. Jupiter (and Saturn) are primarily responsible for this ejection. If star formation were always accompanied by ejection of comets, interstellar space would be well populated by these bodies, which from time to time should penetrate the inner Solar System and be discovered as comets with distinctive hyperbolic orbits. No such comets have been observed, whereas it has been argued that at least one should have been³.

More direct evidence for a paucity of extrasolar Jupiter-like planets is suggested by telescopic searches for perturbations of the motion of nearby stars by possible Jupiter-mass companions^{4–6}. Although more than 20 stars have been well studied, no evidence has been found for companion planets down to a limit of 1 to 3 Jupiter masses at a distance of 5 AU.

If 'Jupiters' are rare, why should we happen to live in an unusual planetary system that contains a Jupiter (and a Saturn as well)? One answer could be that our having a Jupiter is just a random chance. If gas giants can occasionally form somewhere, why not here? A more in-

teresting possibility is that if it were not for the gas giant planets, the Earth would not be habitable, at least for advanced organisms that think about questions of this kind. Maybe we must observe a Jupiter in our Solar System because otherwise we wouldn't be here. Instead of producing Jupiters of roughly 300 Earth masses, perhaps lack of synchronization of the independent processes of planetary growth and gas loss would usually lead to hydrogen-poor 'failed' Jupiters of about the mass of Neptune and Uranus.

In such systems the outer Solar System would not be cleansed of residual cometary material, including bodies the size of Pluto or even larger. Instead, most of these bodies could be stored in nearby trans-Neptunian orbits, from whence they could easily penetrate into the inner Solar System over the entire age of the Solar System⁷. Catastrophic impact events of the kind believed to have caused the mass extinctions at the Cretaceous/Tertiary boundary might then have occurred about every 10⁵ years instead of about once in 10⁸ years. Stabilization of the Earth's obliquity and therefore its climate and habitability by the Moon has been suggested⁸; we may be more dependent on other members of our Solar System than previously thought.

It could justifiably be argued that the theoretically based suggestion that Jupiters are rare is simply an excuse for our ineptitude in producing an adequate theory for the formation of the outer planets. The answer must come from observation. Theory can stitch together threads of observations into a fabric of understanding. The observational data are not yet definitive. Taken altogether, the evidence does raise the possibility that what was thought to be the *sine qua non* of a 'real' planetary system, the presence of planets like Jupiter and Saturn, may actually be unusual. In contrast, more easily made terrestrial planet systems physically similar to ours may be abundant but hazardous unless protected by gas giant planets. □

George W. Wetherill is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road Northwest, Washington DC 20015, USA.

1. Zuckerman, B., Forveille, T. & Kastner, J. H. *Nature* **373**, 494–496 (1995).
2. Koerner, D. W., Sargent, A. I. & Beckwith, S. V. W. *Icarus* **106**, 2–10 (1993).
3. Weissman, P. R. in *Comets in the Post-Halley Era* (eds Newburn, R. L. *et al.*) 463–498 (Kluwer, Dordrecht, 1994).
4. Walker, G. H. *et al.* *Icarus* (in the press).
5. Cochran, W. D. & Hatzes, A. P. in *Planets Around Pulsars* (eds Phillips, J. A. *et al.*) 267–274 (Astr. Soc. Pac. Conf. Vol. 36, 1993).
6. Marcy, G. W. & Butler, R. P. *Publ. astr. Soc. Pacif.* **104**, 270–277 (1992).
7. Wetherill, G. W. in *Planetary Systems: Formation, Evolution and Detection* (eds Burke, B., Rahe, J. H. & Roettger, E. E.) 23–32 (Kluwer, Dordrecht, 1994).
8. Laskar, J. & Robutel, P. N. *Nature* **361**, 608–615 (1993).

110, 111... and counting

Darleane C. Hoffman

THE announcement late last year^{1,2} of the synthesis and detection of the two new elements 110 and 111 by scientists at the Gesellschaft für Schwerionenforschung (GSI) at Darmstadt, Germany, has been received with excitement by scientists around the world. After nearly a decade-long hiatus, these are the first new elements to be discovered since the production and identification of elements 107, 108 and 109 (now known as nielsbohrium, hassium and meitnerium) between 1981 and 1984, also using the electromagnetic velocity filter instrumentation at GSI.

In May 1994, Ghiorso and co-workers at Lawrence Berkeley Laboratory had reported some evidence for the detection of one atom of element 110 with mass 267, produced in bombardments of bismuth with cobalt projectiles in 1991 at the SuperHILAC just before it was shut down. Nonetheless, they laud the GSI group for their superb and definitive experiments.

Early last November, the GSI team of European scientists, led by Peter Armbruster and Sigurd Hofmann, identified¹ the first atom of element 110 with mass 269 in fusion reactions between ²⁰⁸Pb targets and ⁶²Ni ions accelerated in the UNILAC. In subsequent experiments² using ⁶⁴Ni projectiles, they detected six atoms of an isotope of element 110 with mass 271. In additional experiments which ended on 22 December 1994, they identified three atoms of element 111 with mass 272, produced in fusion of stable ²⁰⁹Bi with the ⁶⁴Ni projectiles. The resulting excited compound nucleus ²⁷³111 quickly de-excites by emitting a neutron to form ²⁷²111 which then decays by emission of alpha-particles (see figure).

Unambiguous identification of the mass and atomic number of the new elements was made by observation of their decay via a series of α -particle emissions to known daughter and granddaughter nuclei. In the case of element 110 this included delayed α - α coincidence measurements to the known daughter (and further generation) isotopes of elements 108 to 102, whose decay data had been established in previous experiments. Although more than 3×10^{12} nickel ions per second pass through the targets, the desired reactions are very rare and only four atoms of element ²⁶⁹110 were identified during the two weeks of the experiment. From the measured lifetimes of these atoms, the GSI team calculated the half-life of the new isotope to be only 0.17 ms. However, the heavier isotope, ²⁷¹110, observed in the bombardments with ⁶⁴Ni, had a half-life of 1.4 ms, nearly ten times as long, and also had several times the

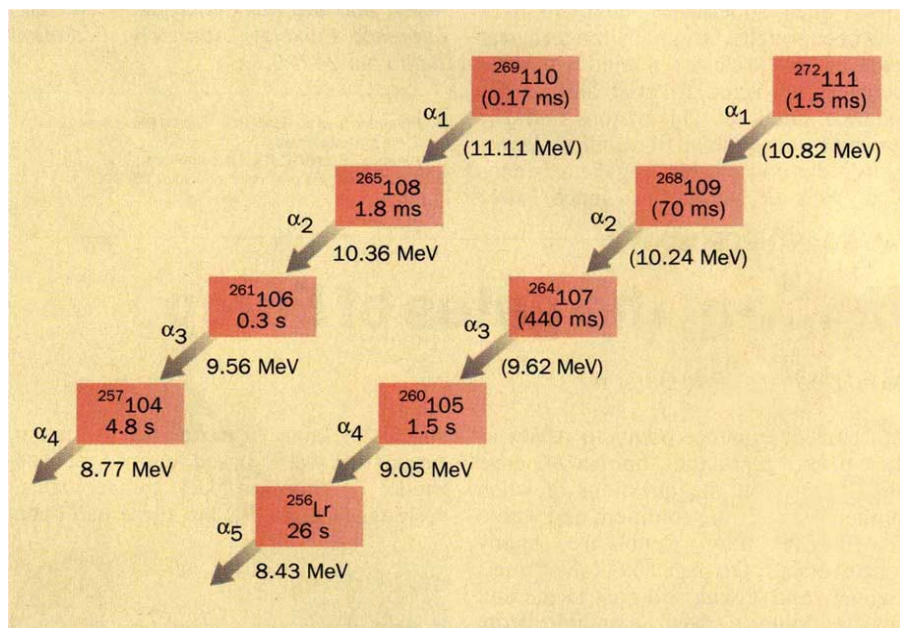
production probability. These results are especially encouraging for the production of still heavier elements and longer-lived isotopes of the known heavy elements — isotopes which may even have long enough lives to permit studies of their chemical properties.

Identification of three atoms of element ²⁷²111, with a half-life of 1.5 ms, was accomplished by the same α - α delayed coincidence technique. This time it involved measuring the α -decay sequence of two previously unknown heavy isotopes of

fore, the GSI team devoted much effort to improving the overall efficiency of the Separator for Heavy Ion reaction Products (SHIP) and the detector system, suppressing the background, using eight targets on a rotating wheel, and increasing the beam intensities. Consequently, it was possible to complete successful experiments over a period of weeks rather than months.

The measurements showed a longer half-life and decreased α -decay energy for the mass 271 isotope of element 110 than for mass 269, consistent with theoretical predictions of increased stability in the proximity of a deformed shell at neutron number 162.

The past year has been a memorable



Decay chains for ²⁶⁹110 and ²⁷²111; the values in parentheses are the half-lives and α -decay energies measured by the GSI group. Lr is lawrencium, element 103, beyond which nomenclature is under dispute.

elements 109 and 107, which ultimately resulted in the production and identification of the known isotopes, ²⁶⁰Ha (105) and ²⁵⁶Lr (103). In this way, positive evidence was provided for the existence of the new isotopes ²⁶⁸109 (70 ms) and ²⁶⁴107 (440 ms), as well as for element 111.

Pre-experiments conducted by the GSI group were essential to these recent discoveries; the object of these experiments was to measure the excitation functions for compound nucleus 'fusion' reactions with lead, in order to determine the optimum projectile energies to be used in the production of the new elements via evaporation of only one neutron. Based on these results, which showed that the maximum cross-sections occurred at sub-barrier projectile energies, cross-sections of the order of a picobarn (10^{-36} cm²) were estimated. Experiments of several months would have been needed if the overall efficiency had been as low as it was in the discovery of elements 107 to 109. There-

one for heavy element scientists. Earlier in 1994, a collaboration of scientists from Lawrence Livermore National Laboratory in the United States and from the Joint Institutes for Nuclear Research at Dubna, Russia, reported evidence for mass 265 and 266 isotopes of element 106 (named seaborgium by its discoverers) with half-lives inferred to be as long as 2 to 30 seconds, much longer than its previously known isotopes. These results helped to corroborate predictions by theoretician Adam Sobiczewski³ and his group in Warsaw of the existence of a region of enhanced stability around 108 protons and 162 neutrons due to the effect of deformed neutron and proton subshells. The Dubna-Livermore collaboration also conducted experiments in 1994 to produce element 110 with mass 273 in the 'hot-fusion' reaction of ³⁴S projectiles with targets of ²⁴⁴Pu, the most stable isotope of plutonium. In such reactions, the fusion product is highly excited and must emit

five neutrons to produce the desired product. They have identified one decay chain attributable to $^{273}\text{110}$, but are still analysing their data.

All of these developments have caused much optimistic activity centring around the prospects of producing still heavier elements and longer-lived isotopes of currently known elements. Nuclear chemists hope that some of the new isotopes may live long enough to permit chemical studies which can now be conducted on the timescale of seconds. An international group is preparing to do such studies on element 106 at GSI in the next year.

Plans for newer, still more efficient instrumentation for the detection of elements 112, 113 and 114 are proceeding apace. Such 'superheavy' elements have long been sought. Originally the term was reserved for the cluster of relatively stable elements predicted to exist around the spherical shells at 114 protons and 184 neutrons — the 'island of stability'. Investigators are now establishing the existence of a 'rock of stability' at much lower

neutron numbers. The isotopes $^{271}\text{110}$ and $^{272}\text{111}$, as well as some of the longer-lived isotopes of elements 106 and 108, may be stabilized by their proximity to the deformed shells at 162 neutrons and 108 protons, and it has been proposed that these qualify as 'superheavy' elements. But detection of 'true' island elements requires creation of nuclides with neutron numbers near 184 — some 20 neutrons more than the most neutron-rich isotope known to date. Nevertheless, the current rock will be a stepping stone to allow us to reach element 114 even before reaching the spherical closed shell at 184 neutrons.

Darlene C. Hoffman is in the Department of Chemistry, University of California, Berkeley, and the Nuclear Science Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA.

1. Hofmann, S. *et al.* *GSI-94-82 Preprint* (November 1994); *Z. Phys. A* (submitted).
2. Hofmann, S. *et al.* *Z. Phys. A* (submitted).
3. Sobczewski, A. *Phys. Particles Nuclei* **25**, 295 (1993).

questioned both as to the age and human origin of the purported artefacts².

Then Swisher *et al.*⁴ reported high-precision dates of 1.6–1.8 Myr for the oldest fossils at the two Indonesian sites, reopening the questions of whether these Asian populations indeed represented a different species from the African and whether mode 2 (Acheulean) archaeological industries had developed by the time hominids left Africa. The earliest African Acheulean is now dated at about 1.4 Myr in Ethiopia¹⁰, and similar assemblages are known from 'Ubeidiya, Israel, estimated to be of equivalent age (but unassociated with diagnostic human remains)¹¹.

Gabunia and Vekua³ now shift attention westwards again to Dmanisi, which was a major centre between the eighth and twelfth centuries. Excavations there revealed a number of deep pits for grain storage dug into the floors of now-collapsed buildings. In 1981, the walls of such a pit revealed animal bones *in situ* which were identified as belonging to species of Late Villafranchian (Late Pliocene to Early Pleistocene) age, according to the well-documented sequence of mammalian faunas of the Black Sea region. Further work resulted in the recovery of mode 1 artefacts and the development of a stratigraphic profile with normal geomagnetic polarity throughout and a basal basalt dated to about 1.8 Myr. Finally, in 1991, the team found a well preserved human mandible in one of the storage pits.

PALAEOANTHROPOLOGY

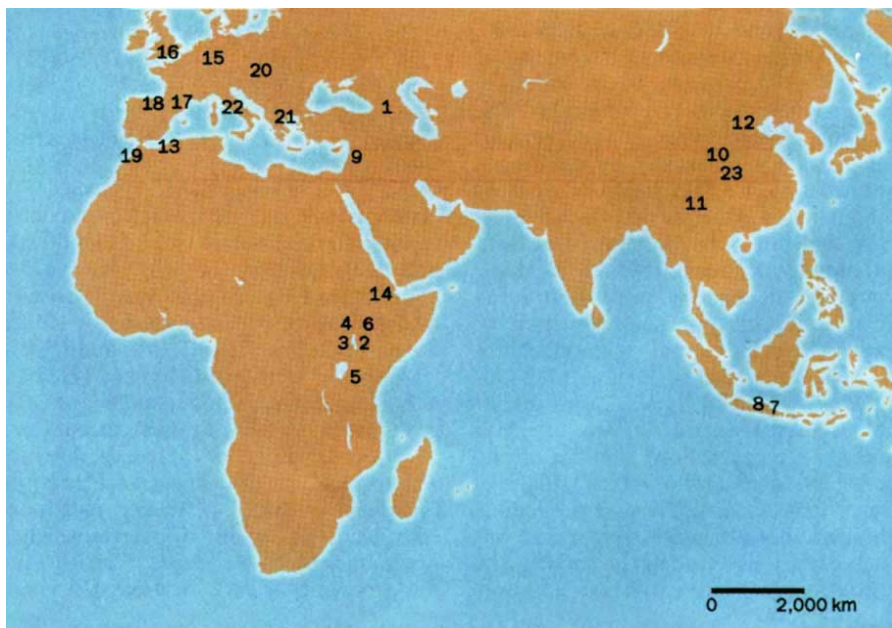
Homo at the gates of Europe

David Dean and Eric Delson

ALL current evidence points to Africa as the place that the human lineage originated¹. But the questions of when humans first left the continent and where in Eurasia they went are highly controversial². On page 509 of this issue³, Gabunia and Vekua add grist to the mill by describing a fossil mandible from Dmanisi, Georgia, attributed to *Homo erectus*. The mandible has been dated by three independent methods to about 1.8 million years old (Myr), making it comparable in age to the most ancient specimens from two sites in Java⁴; but although the dates provided by Swisher *et al.*⁴ are technically excellent, whether or not they apply to the human fossils at both sites is still questionable^{5,6}.

Until a year ago, most researchers would have accepted a date of about 1.25 Myr as a maximum for the first appearance of *Homo erectus* in Eurasia, either in China (Lantian Gongwangling or Yuanmou) or Java (Sangiran or Mojokerto). Older ages had been proposed, but they were mainly discounted, and even the 1.25 Myr Rubicon was questioned as possibly too old. In Africa, the first appearance of this species, there sometimes termed *Homo ergaster*, was dated to about 1.8–1.9 Myr⁷. The earliest human fossils in western Eurasia were dated to 0.5 Myr or less, and most palaeoanthropologists identified these as representing not *H. erectus* but early populations of *H. sapiens* or perhaps a distinct intermediate

species⁸. Claims of more ancient human occupation were based upon supposed mode 1 (Oldowan-like) archaeological residues (see ref. 9), but these had been



Fossil localities mentioned in the text and other important sites of early *Homo*. 1, Dmanisi; 2, Koobi Fora (East Turkana); 3, West Turkana (Nariokotome); 4, Omo Shungura; 5, Olduvai Gorge and Lake Ndutu; 6, Konso-Gardula; 7, Mojokerto; 8, Sangiran; 9, 'Ubeidiya; 10, Lantian (Gongwangling); 11, Yuanmou; 12, Zhoukoudian; 13, Tighenif (Ternifine); 14, Bodo; 15, Mauer (Heidelberg); 16, Boxgrove; 17, Arago; 18, Atapuerca; 19, Thomas Quarries; 20, Vértesszöllös; 21, Petralona; 22, Fontana Ranuccio; 23, Yunxian. Sites 1–8 and 10–12 have yielded *Homo erectus*; site 9 early artefacts and unidentified human bone fragments; sites 5 and 13–23 early 'archaic' *Homo sapiens*.

Features of the Dmanisi mandible

THE molar teeth of the jaw show a great size reduction from front to back, greater than that known for any Asian (or African) *H. erectus* fossil or sample, and similar in some ways to that seen in later 'archaic' specimens of *H. sapiens*, such as those from Atapuerca (Spain). The rami are broken away, but they would almost certainly have splayed out postero-inferiorly as in African and Chinese *H. erectus*; moreover, the rear of the tooth row is crowded, as evidenced by the reduced M_3 and lack of a retromolar space. As in all *H. erectus*, the symphysis in the Dmanisi mandible is rounded anteriorly, although it does not slope back as steeply as in some early African individuals (for example WT 15000; ref. 14). On the other hand, the Dmanisi jaw shows no sign of the lateral tubercles on either side of the symphysis which are responsible for the mental trigone in the earliest 'archaic' *H. sapiens*. These tubercles are probably the result of greater

'wishboning' stress due to increased brachycephaly and klinorhynchism in early *H. sapiens*. The angulation of the mental area is less important than this biomechanical shift, which should be the subject of renewed investigation.

All the mandibular differences may reflect evolution towards a wider cranial base and perhaps other 'gestalt' changes in earliest 'archaic' *H. sapiens* as compared to *H. erectus*. Increased brachycephaly would result in less prognathism and greater distance across the mandibular condyles, so that the gonial area need not splay out to provide proper position for the masseter muscle, and the symphysis, in effect, would 'migrate' back to the lateral tubercles as it widens. As the anterior dentition later becomes smaller, the intertubercular width decreases and the general gracility leads to the modern trigone and mental eminence.

This new interpretation of the anterior mandible leads to a better understanding of

the value of this region in taxonomic evaluation of individual fossils — for example, the Tighenif (ex-Ternifine) mandibles appear to display lateral tubercles, which led one of us (D.D.) to transfer that population to *H. sapiens*¹⁸. In turn, given the estimated age of 0.6–0.7 Myr for Tighenif and the age of ~0.6 Myr for the Bodo (Ethiopia)¹⁹ cranium of early *H. sapiens*, it is likely that, at the least, earliest 'archaic' *Homo sapiens* populations were present in the northern half of Africa before the earliest known anatomical evidence for their presence in Europe, ~0.5 Myr². Recent evidence from East Asia suggests that there was a late but long (0.3–0.15 Myr) period of overlap between 'archaic' *H. sapiens* and late-surviving *H. erectus*. This implies that the routes to East Asia might not have been open to 'archaic' *H. sapiens* until after Europe had been successfully inhabited.

D.D. & E.D.

Analytical details have not been published for either the date or the palaeomagnetic profile, but they are concordant with a correlation to the Olduvai normal geomagnetic subchron, currently dated between 1.77 and 1.95 Myr⁶; the base of the Pleistocene is defined as just younger than the end of this interval. This would fit closely with the estimated age of the faunal assemblage, which might date at between 2.1 and 1.4 Myr. If indeed the polarity is normal, no other option remains: the Réunion Subchron(s) date to 2 Myr or older¹², while the only younger normal subchrons are the newly revitalized Cobb Mountain at about 1.15–1.2 Myr¹³ and the Jaramillo spanning 1.0 Myr. The associated fossil mammals, including rodents of Pliocene affinity and no species of Biharian age, count as strong evidence against the younger dates. On the other hand, it is conceivable that the palaeomagnetic analysis did not involve sufficient (alternating field) demagnetization to remove later overprinting, in which case the fossils *could* be of somewhat younger age.

There is little doubt that the mandible itself can be attributed to *Homo erectus*. We do not consider that the African Pleistocene fossils sometimes termed *H. ergaster* represent a distinct biological species, given the known ranges of variation — ref. 7 supports the concept of *H. ergaster*, while refs 14 and 15 reject it (as we do). But several features of the Dmanisi jaw are quite distinctive, especially when taken in combination. The technical details and their implications are discussed in the boxed text.

As Gabunia and Vekua note, if the Dmanisi specimen indeed dates to 2.0–1.5 Myr, roughly 1 Myr passed between then and the first widespread occurrence of

humans farther west, in Europe proper. Was this time gap the result of climatic severity, such that pre-*sapiens* humans were unable to survive long in the periglacial environments of even southern Europe? Or was it related to their inability to cope with the large terrestrial carnivores, as hypothesized by Turner¹⁶? Further studies of the archaeological residues at Dmanisi, as well as determination of whether the fauna was human food refuse or the result of other taphonomic circumstances, should help to answer these questions. As Bar-Yosef¹⁷ points out, most of the mammalian fauna of Pleistocene southwest Asia is derived from farther east or perhaps Europe; little of it, other than *Homo*, comes from Africa. Whatever the sequence of sites, *H. erectus* must have traversed southwest Asia en route to Indonesia, China and the rest of Eurasia, but various attempts may have been made before one or more colonizations were successful¹⁷.

Perhaps the situation in the Early Pleistocene was analogous to that around 0.1 Myr, when (according to one view) early modern humans entered southwest Asia from Africa but did not move into Europe for over 50,000 years. It may be that

cultural adaptation, involving toolkit technology and perhaps broader social evolution, was at the root of both successful invasions of western Eurasia, by 'archaic' and anatomically modern *Homo sapiens*, respectively. Given our growing understanding of *Homo erectus* and 'archaic' *Homo sapiens* palaeodemography and migration patterns, it is becoming ever more clear that these two species form an ancestor–descendant pair in the Middle Pleistocene, with the latter replacing the former from west to east. We are now faced with the continued challenge of finding not only more evidence of these extinct populations, but also the palaeoclimatic and perhaps technological conditions which allowed or prevented them from occupying various regions of the Old World, including the cul-de-sac we know as Europe. □

David Dean is in the Department of Anatomy, Case Western Reserve University, Cleveland, Ohio 44106-4905, USA. Eric Delson is in the Department of Anthropology, Lehman College, and the Graduate School, City University of New York, Bronx, New York 10468, USA, and at the American Museum of Natural History.

- White, T. D., Suwa, G. & Asfaw, B. *Nature* **371**, 306–312 (1994).
- Roebroeks, W. & van Kolfschoten, T. *Antiquity* **68**, 489–523 (1994).
- Gabunia, L. & Vekua, A. *Nature* **373**, 509–512 (1995).
- Swisher, C. C. et al. *Science* **263**, 1118–1121 (1994).
- De Vos, J. & Sondaar, P. *Science* **266**, 1726–1727 (1994).
- Swisher, C. C. *Science* **266**, 1727 (1994).
- Wood, B. A. *Courier Forschungs-Institut Senckenberg* **171**, 159–165 (1994).
- Rightmire, G. P. *The Evolution of Homo erectus* (Cambridge Univ. Press, 1990).
- Bonifay, E. & Vandermeersch, B. (eds) *Les Premiers Européens* (Comité Travaux Historiques Scientifiques, Paris, 1991).
- Asfaw, B. et al. *Nature* **360**, 732–735 (1992).

- Verosub, K. L. & Tchervov, E. in *Les Premiers Européens* (eds Bonifay, E. & Vandermeersch, B.) 237–242 (Comité Travaux Historiques Scientifiques, Paris, 1991).
- McDougall, I., Brown, F. H., Cerling, T. E. & Hillhouse, J. W. *Geophys. Res. Lett.* **19**, 2349–2352 (1992).
- Renne, P. R. et al. *Geology* **22**, 783–786 (1994).
- Walker, A. & Leakey, R. (eds) *The Nariokotome Homo erectus Skeleton* (Harvard Univ. Press, Cambridge, MA, 1993).
- Brauer, G. *Courier Forschungs-Institut Senckenberg* **171**, 301–318 (1994).
- Turner, A. *Courier Forschungs-Institut Senckenberg* **171**, 241–247 (1994).
- Bar-Yosef, O. in *Origins of Anatomically Modern Humans* (eds Nitecki, M. H. & Nitecki, D.) 23–66 (Plenum, New York, 1994).
- Dean, D. thesis (City Univ. New York, 1993).
- Clark, J. D. et al. *Science* **264**, 1907–1910 (1994).

Understanding ocean dynamics

Kurt Lambeck

SATELLITE altimetry has been one of the cornerstones of geodynamics research from space for a decade and a half, but the oceans have had to wait longer for their turn. A meeting in December* included symposia dedicated to the scientific results of two particularly promising missions: Topex/Poseidon, jointly run by the United States and France, and the European Space Agency ERS-1. What emerged, as illustrated here, was an exciting, if preliminary, insight into the dynamics of the world's oceans.

The satellite-borne altimeter measures the spacecraft height above the ocean surface. If the orbit of the spacecraft is accurately known then the shape of the ocean can be directly mapped. To a first approximation this shape is the geoid, an equipotential surface defined by the Earth's internal mass distribution, so the altimeter observations have provided valuable information on the physics of the solid Earth. Of greater interest to the oceanographic community is the departure of the sea surface from the geoid, the 'sea surface topography'. This topography and its changes provide the important information on the ocean currents. Earlier altimeter missions were hampered by either inadequate orbit control or a short spacecraft lifetime and consequently not much was learnt about the oceans. With the initial results from the Topex/Poseidon and ERS-1 missions, this has changed dramatically. The cause of this change has been progress in orbital dynamics and tracking accuracies which allow radial orbital accuracies to be maintained — and temporal variations in the ocean surface to be measured — to an accuracy of a few centimetres.

What is being revealed by the satellites is the ocean at work in real time. Much of what is seen has been known in outline, inferred from sparse shipboard measurements, but the full dynamic nature of the ocean is now emerging with unprecedented resolution. The tides are readily seen, not only the main diurnal and semi-diurnal components but also the longer-period tides such that at any time the global ocean tide is known to within 2 or 3 centimetres accuracy, a huge improvement over earlier results. It is encouraging that numerical models for the tides are also much improved and surely we can expect to see, for example, the definitive study of the mechanisms and distribution of the dissipation of tidal energy in the oceans. Intense, small-scale motions portray a complex instantaneous flow field

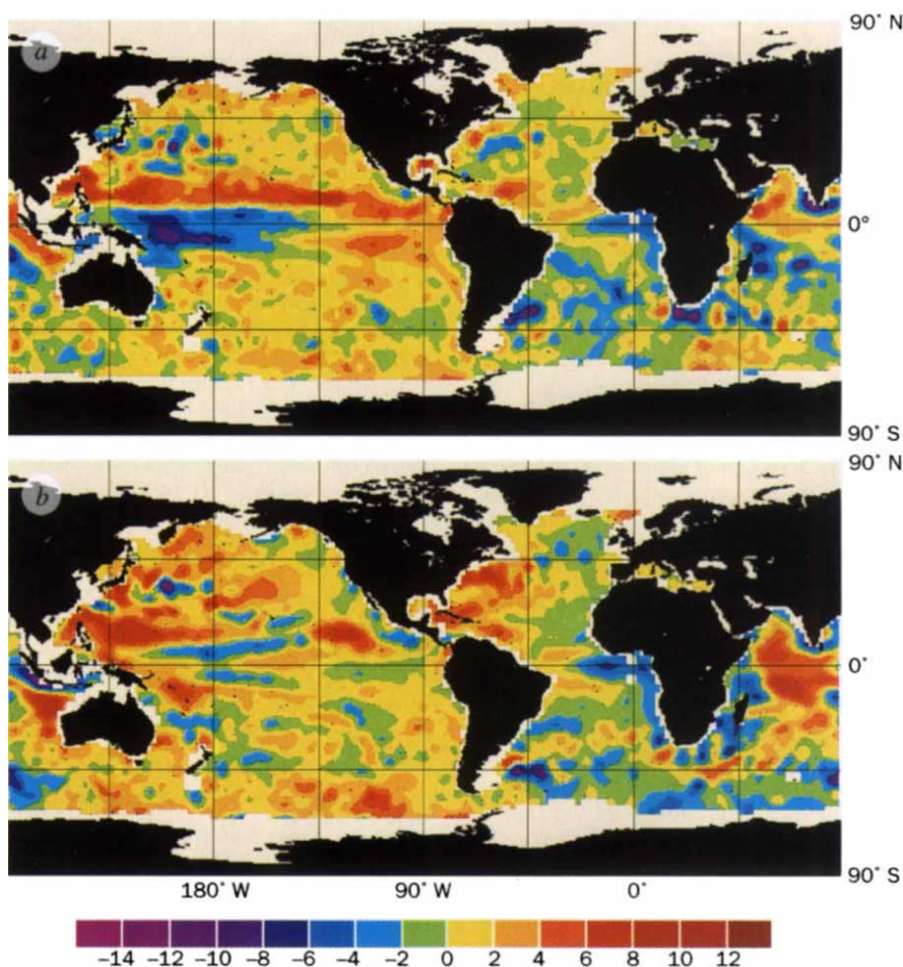
which tends to swamp the weaker but larger-scale motions that are of greater significance to climate. Early results are indicating that the latter can be extracted from the higher-frequency signals.

Seasonal cycles, for example, are being mapped by the Topex/Poseidon team. Perhaps most striking of these results is the asymmetry between the two hemispheres. The highest and lowest sea levels occur in the Northern Hemisphere autumn and spring respectively, with amplitudes that are much greater than for the Southern Hemisphere. Perhaps, then, the heat storage in the smaller Northern Hemisphere ocean is much greater than in the Southern Hemisphere. The seasonal pattern has now been established for the years 1992–94 but the interannual variability appears almost as great as the seasonal change itself. For example, there are notable differences in the Pacific equatorial region for the summer and winter months which presumably

relate to the El Niño/Southern Oscillation. This can be seen in the figure for sea surface heights in the summers of 1993 (a) and 1994 (b). But note also the differences for these two records in the western North Atlantic and in the Indian Ocean.

The complementarity of the two altimeter missions is important. The orbital configuration of the Topex mission is such that it provides mainly a global-scale picture of the oceans and the high spatial and temporal variability is less well observed. ERS-1, whose orbit is such that the satellite flies over the same part of the ocean more frequently, provides more information on such mesoscale signals and processes. At the moment, it seems that some orbital problems remain to be ironed out and the full potential of the mission is yet to be achieved. Post-event processing is improving matters, and much new information can be expected soon from the smaller but important shallow areas of the oceans.

Determination of the very long-term variations in the ocean dynamics remains a matter for the future. First, long time series of data are required. Topex/



Sea surface topography for the Northern Hemisphere summers of (a) 1993 and (b) 1994. Each result represents a three-month mean of the sea surface elevations with respect to the mean surface established for the two-year period. The scale is in centimetres. Results are from the Topex/Poseidon Science Team led by C. Wunsch, J.-F. Minster and C. Koblinsky.

*Fall Meeting of the American Geophysical Union, San Francisco, 6–9 December 1994.

Poseidon is expected to live for at least three more years, and the space agencies should recognize the importance of providing continuity in order to study the longer-term interannual variability such as the Southern Oscillation and El Niño phenomena. Second, there remains a need for better information on the shape of the geoid. At present, this surface is less well-known than the measurement accuracy of the altimeter or the accuracy of the tracking data and it remains impossible to measure the absolute shape of

the sea at the requisite level of accuracy. A dedicated mission to determine the planet's external gravity field, coupled with future altimeter missions like Topex and ERS, could revolutionize our understanding of the dynamics of the oceans and of their interaction with the atmosphere, as well as contributing to the deciphering of the Earth's sub-surface structure. □

Kurt Lambeck is in the Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia.

MATERIALS SCIENCE

Multiple Kauzmann paradoxes

Robert W. Cahn

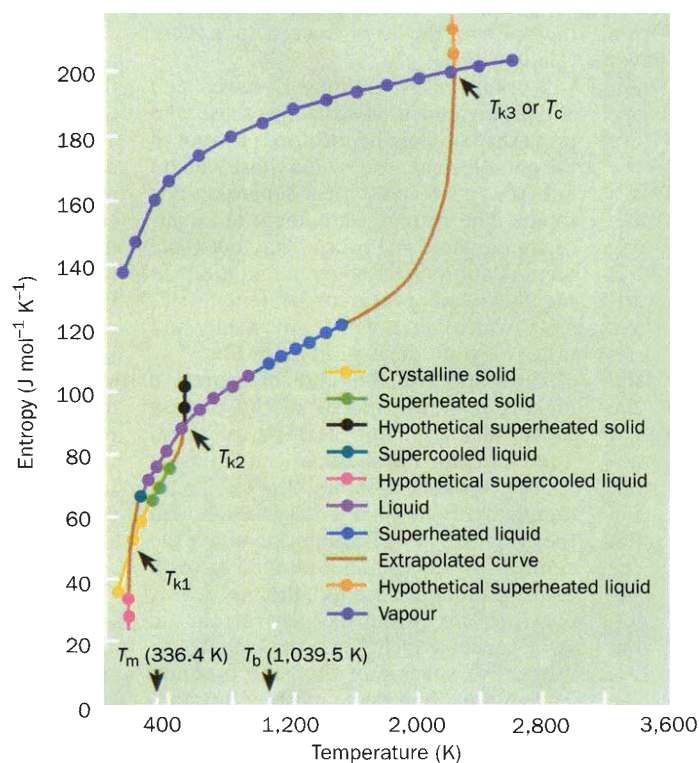
THE Kauzmann paradox, beloved of physical chemists (and chemical physicists) has featured in these pages before¹. The concept has been gradually broadened, and a recent paper from Bangalore, India, in the *Journal of Chemical Physics*², has broadened it further still.

Kauzmann's original paper³ first drew attention to the fact that the entropy difference between the crystalline phase and the (extrapolated) liquid phase of a substance can vanish at a temperature well above absolute zero; below that temperature, the crystal would have a higher entropy than the liquid. As this would violate the Third Law of Thermodynamics, Kauzmann termed his conclusion a paradox. A good concise introduction to this paradox appears in Elliott's book on amorphous solids⁴, and a shorter one in my News and Views piece in 1988. The usual escape from the paradox is to claim that the 'ideal glassy state', which is the equilibrium configuration of the liquid at the Kauzmann temperature, can never be attained because the diffusion rate becomes negligible before the temperature has sunk to that level. In other words, the amount of supercooling a liquid can accept while remaining in metastable equilibrium is limited.

Fecht and Johnson⁵ and Fecht⁶ extended Kauzmann's ideas to high temperatures: the question they examined is whether a crystal can be superheated to (or beyond) the temperature at which the liquid and the superheated (extrapolated) crystalline phase have the same entropy.

Their arguments lean heavily on the very high vacancy concentrations that are to be expected at this high 'isentropic temperature'; the crystal collapses at this temperature because the vacancies destabilize the lattice. Such a collapse is termed an 'entropy catastrophe'.

Kishore and Shobha² now raise the stakes by considering the relationship be-



tween a superheated liquid and its vapour. The figure, for potassium, one of the specific examples treated in their paper, shows the entropies of solid, liquid (shading into glass at low temperatures) and vapour. T_{k1} and T_{k2} are the original Kauzmann temperature and the upper Kauzmann temperature as previously analysed by Fecht and Johnson. T_{k1} is the absolute

lower temperature limit for the existence of a supercooled liquid (*alias* glass) rather than the crystalline solid. T_{k2} is the absolute upper temperature limit for the superheated crystalline solid, vis-à-vis the liquid. The innovation is T_{k3} , the absolute upper temperature limit for a superheated liquid vis-à-vis the vapour phase. Kishore and Shobha explicitly remark that T_{k3} "could be considered as the critical temperature, T_c ".

They compute values of T_{k3} for water and for various alkali and other metals. Agreement is mostly only moderate, within about 15 per cent in terms of absolute temperatures. For the example of potassium in the figure, $(T_{k3})_{calc} = 1,932$ K whereas $(T_c)_{exp} = 2,220$ K. For water, agreement is much better: $T_{k3} = 636$ K, $T_c = 643$ K. No attempt is made in the paper to bring in pressure as a variable, and it would have been better to have attempted a calculation of T_{k3} at the critical pressure, for comparison with T_c . In spite of this reservation, the ghost of Van der Waals will be pleased with this new approach to his famous criticality concept that goes back all the way to his 1873 doctoral thesis (which, as Maxwell remarked at the time, "at once put his name among the foremost in science").

As John Cahn has reminded us⁷, it was van der Waals who first recognized that a liquid at the critical point has negative compressibility which renders it unstable, and it was he who coined the word 'spinodal' for this condition: after a gap of almost a century, the relevance of the spinodal concept for certain solid-state transitions, now known as spinodal transformations, came to be perceived. I wonder whether, one day, spinodal transformations and the Kauzmann paradox will be linked up.

Quite generally, Kishore and Shobha find that $2T_{k1} \approx T_m \approx T_{k2}/2$, where T_m is the melting temperature. They go on to estimate T_{k1} and T_{k2} by an alternative approach: they advance arguments as to why the configurational part of entropy is much more important here than the vibrational part, and use indirect (nonthermal) input, for instance viscosity measurements via the well-known Vogel-Tammann-Fulcher equation for the viscosity of liquids (which is focused on configurational features), to estimate configurational entropies and hence the Kauzmann temperatures. They achieve reasonable agreement with values

deduced from thermal inputs. This approach also yields a 'figure of merit', characteristic for whole groups of substances (for example, the alkali metals) which can serve as a measure of what they term the 'fragility' of the group of substances — a curious concept, going back some years⁸, which expresses how robust a phase is. With respect to a liquid, a high 'fragility' expresses easy glass formation (a high glass transition temperature), or a strong 'glass-forming temperament', another novel phrase introduced here.

The authors' ideas are explored not only for pure metals but also for a Na-K alloy; this is clearly just a start in applying Kauzmann concepts more widely to alloys. Briefly, they also discuss polymers. Now that we have a complete gamut of

three Kauzmann temperatures, it is to be hoped that the relationships between them will be further explored, theoretically and experimentally, and more particularly, that the connection between T_{k3} and T_c will be investigated rigorously. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

1. Cahn, R. W. *Nature* **334**, 17–18 (1988); **342**, 619–620 (1989); **356**, 108–109 (1992).
2. Kishore, K. & Shobha, H. K. *J. chem. Phys.* **101**, 7037–7047 (1994).
3. Kauzmann, W. *Chem. Rev.* **43**, 219–256 (1948).
4. Elliott, S. R. *Physics of Amorphous Materials*, 29–30 (Longman, London, 1984).
5. Fecht, H. J. & Johnson, W. L. *Nature* **334**, 50–51 (1988).
6. Fecht, H. J. *Nature* **368**, 133–135 (1994).
7. Cahn, J. W. *Scand. J. Metall.* **20**, 9–17 (1991).
8. Angell, C. A. *J. non-cryst. Solids* **73**, 1 (1985).

ALZHEIMER'S DISEASE

Mouse model made

Karen Duff and John Hardy

MANY position papers¹ on Alzheimer's disease (AD) point out that progress in studying the disorder is hampered by the lack of an animal model. Creating such a model has been no easy task. However Games and colleagues, who describe their work on page 523 of this issue², have produced transgenic mice that will go a long way to answering the needs of those investigating how the disease develops.

After heart disease, cancer and stroke, Alzheimer's disease is the fourth leading cause of death in the developed world. It is a disease which afflicts a very high proportion of the elderly. The symptoms begin with problems in forming new memories but, as the disease progresses, more and more higher functions are lost and the sufferer usually dies after several years of illness, demented, bedridden and incontinent. Because of this lengthy progression, it is possibly a disease without equal in terms of the financial and emotional toll it puts on those who have to care for the victim.

The pathology of the disease is complex, but there are two hallmarks that occur abundantly and with a characteristic distribution in most cases. These are the neuritic plaque, a largely extracellular lesion consisting of β -amyloid with other components^{3,4}, and the neurofibrillary tangle, consisting largely of abnormally phosphorylated tau protein⁵. Besides this frank pathology, neuronal and synaptic loss and reactive gliosis also occur. Unlike heart disease, cancer and stroke, there has been no animal model of the disease and Alzheimer's pathology has only been seen in humans (apart from very sparse examples in the occasional elderly primate, dog or sheep). A good model could help

considerably, both in understanding the relationship of the various aspects of the pathology to each other and in testing therapies for the disease based upon these relationships.

Workers investigating AD have been sharply divided into two camps: those who think that amyloid deposition is central to the pathological process and those who do not (see refs 6 and 7 for a debate on this topic). The strongest argument in favour of the amyloid hypothesis has been the identification of pathogenic mutations in the β -amyloid precursor protein (APP) gene which alter APP processing and make amyloid deposition more likely^{8,9}.

The work of Games and colleagues² in making a mouse model of the disease settles this argument, perhaps for good. The authors have taken an unusual construct which includes the full-length human APP complementary DNA with the APP717Val→Phe mutation under the control of the platelet-derived growth factor promoter. Inserted into this cDNA construct are introns 6, 7 and 8 from the APP gene which are important for alternative splicing of the gene product. The authors then made mice by the standard microinjection route (for a review of other attempts at transgenics for AD, see ref. 10). Probably a key reason for the success in generating pathology in these mice is that this construct (which is a cross between a cDNA and a genomic construct) directs much higher levels of expression than standard cDNA constructs. The level of APP expression is probably also greater than that in mice bearing the APP yeast artificial chromosome (YAC) constructs. If it is, this may underlie the presence of Alzheimer-like features in

these mice and their absence in the YAC APP mice^{10–12}.

The pathology of the mice is striking. As in AD, there are neuritic plaques and neuronal and synaptic loss with reactive gliosis. This pathology appears to have broadly the same neuroanatomical distribution as in AD, cortical and hippocampal pathology being most severe. But, as yet, there is no evidence for the occurrence of neurofibrillary tangles, despite the presence of swollen and dystrophic neurites, and the mice have not yet been assessed for cognitive dysfunction.

Although there is no doubt that much of the pathology resembles that of AD, there will now be an enormous amount of detailed comparison of the pathology in these mice with that in humans. It is likely that those who are sceptical of the amyloid-cascade hypothesis will draw comfort from the apparent absence of tangles in the mice. However, it remains possible that their absence reflects the fact that they are merely a marker of cell injury in some mammals that does not occur in others: indeed, in some cases of dementia, neuritic plaques occur without tangles¹³, meaning that tangles may indeed be epiphenomena as they are in some forms of prion disease¹⁴.

The production of these remarkable animals raises three types of issue. First, there are the practical matters to do with the production of the mice themselves. What element of the construct was crucial to its success (the promoter, the use of introns or the use of mutations)? Will mice containing one or more pathogenic mutations carried on full genomic constructs (for example in YACs) be even better^{11,12,15}? Second, there are a huge number of questions relating to the pathogenesis of the disease. For instance, can the disease process in these mice be modified by breeding them with mice containing different human apolipoprotein E genes as the human disease is^{16,17}? If the disease spreads down neuronal pathways¹⁸, will it be possible to alter the

1. *Neurobiol. Aging Suppl.* **2**, 15 (1994).
2. Games, D. et al. *Nature* **373**, 523–527 (1995).
3. Glenner, G. G. & Wong, C. W. *Biochem. Biophys. Res. Commun.* **120**, 885–890 (1984).
4. Masters, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4245–4249 (1993).
5. Lee, V. et al. *Science* **251**, 675–678 (1991).
6. Selkoe, D. J. *J. Neuropath. exp. Neurol.* **53**, 438–447 (1994).
7. Roses, A. D. *J. Neuropath. exp. Neurol.* **53**, 429–437 (1994).
8. Goate, A. et al. *Nature* **349**, 704–706 (1991).
9. Suzuki, N. et al. *Science* **264**, 1336–1340 (1994).
10. Lamb, B. T. *Nature Genet.* **9**, 4–6 (1995).
11. Lamb, B. T. et al. *Nature Genet.* **5**, 22–30 (1993).
12. Pearson, B. E. & Chou, T. K. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10578–10582 (1993).
13. Hanson, L. et al. *Neurology* **40**, 1–8 (1990).
14. Dlouhy, S. R. et al. *Nature Genet.* **1**, 64–67 (1992).
15. Duff, K., McGuigan, A., Huxley, C., Schulz, F. & Hardy, J. *Gene Therapy* **1**, 71–75 (1994).
16. Corder, E. H. et al. *Science* **261**, 921–923 (1993).
17. Houlden, H. et al. *Lancet* **342**, 737–738 (1993).
18. Pearson, R. C. A. & Powell, T. P. S. *Rev. Neurosci.* **2**, 101–122 (1989).

disease process in the mice by selective lesioning? And does the neurochemistry of the deficit in the mouse resemble that of AD? Finally — and perhaps of most importance in the long run — it remains to be seen whether the mice will be useful for testing therapeutic compounds.

The answers to many of these questions will now come quickly, especially if these mice are made available to the general

research community. Clearly, expanding and distributing colonies of the mice should be a priority. □

Karen Duff and John Hardy are in the Suncoast Alzheimer's Disease Laboratories, Departments of Psychiatry, Pharmacology, Neurology and Biochemistry, University of South Florida, Tampa, Florida 33613, USA.

PROTEIN-TYROSINE KINASES

Getting down to specifics

Tony Pawson

PROTEIN kinases, by definition, work by phosphorylating specific substrates. Identifying the element that is recognized by any protein kinase is a crucial issue, as it provides a key to defining its biological targets. On page 536 of this issue¹, Songyang and co-workers summarize a technique for rapidly assessing the substrate specificities of protein kinases, and apply this procedure to an analysis of protein-tyrosine kinases. The results are intriguing.

Protein kinases vary widely in their affections for protein substrates. Some are essentially monogamous, and go through life phosphorylating one or only a few substrates. The Map kinase kinase (or Mek) is a good example, in that it specifically phosphorylates Map kinase on both threonine and tyrosine residues, and requires an intact Map kinase protein as a substrate². Other protein kinases bestow their favours liberally, apparently because they recognize short peptide motifs, including the hydroxyamino acid to be phosphorylated and a variable number of flanking residues³. Hence Map kinase, unlike its activator Mek, has a large number of known and potential substrates, which it phosphorylates at a serine or threonine followed by a proline⁴. Identifying *in vivo* substrates for protein kinases, and assessing the determinants recognized by their catalytic domains, has always been a herculean task. Songyang and colleagues have now come to our rescue with a clever yet simple approach to the problem.

The method they used was described recently⁵. Prompted by their previous success in defining the binding specificities of Src-homology-2 (SH2) domains using a degenerate phosphopeptide library⁶, the authors have adapted this approach to examine protein kinase specificity. In the original procedure, designed to look at serine/threonine-specific protein kinases, a degenerate peptide library was constructed with the sequence shown in Fig.

1; the central serine is the potential phosphorylation site, and other fixed residues allow for recovery of phosphorylated peptides, or their quantitation. The idea, which is simple enough, is to incubate the library with a purified protein kinase, isolate the phosphorylated products and sequence them. If a specific amino acid is enriched at one of the degenerate (X) positions in the phosphopeptide, compared with the unphosphorylated starting material, this indicates a selection by the protein kinase and implies that the residue

Met-Ala-X-X-X-X-Ser-X-X-X-X-Ala-Lys-Lys-Lys

FIG. 1. Degenerate peptide sequence used by Songyang *et al.* to investigate the specificity of serine/threonine protein kinases. In their latest work, discussed here, the central Ser was replaced by Tyr. X represents any amino acid except Trp, Cys, or the relevant hydroxyamino acid.

concerned forms part of the substrate recognition site.

In their original paper, Songyang *et al.* employed kinases with well-defined substrate specificities. In the peptide library, the cyclic-AMP-dependent protein kinase showed a strong selection for arginine at the positions three and two residues amino-terminal to the serine (the -3 and -2 positions) and for a hydrophobic residue at the residue immediately carboxy-terminal to the serine (the +1 position). This conforms remarkably well to the known specificity of this kinase. Similar success was achieved with cyclin-dependent kinases.

On to protein-tyrosine kinases (PTKs)! Transmembrane receptors with PTK domains are prone to autophosphorylate upon activation, and consequently to bind the SH2 domains of downstream signalling proteins⁷. SH2 domains are small modules, found in a diverse group of intracellular proteins, that recognize specific phosphotyrosine-containing motifs. Notably, SH2 domains contact at least three residues carboxy-terminal to the phosphotyrosine. Many PTKs, unlike the receptors, are entirely confined to the inside of the cell, where they can form

signalling subunits of multi-chain receptors, such as those for antigens or cytokines. Curiously, the Src, Abl and Fps/Fes intracellular kinases (and their relatives) each have an intrinsic SH2 domain located adjacent to the catalytic domain. Indeed the SH2 domain was originally identified as an element that contributes to the activity and substrate recognition of the Fps/Fes PTK (ref. 8), and has been shown to modify the spectrum of phosphotyrosine-containing proteins induced by the Abl or Src kinase domains *in vivo*⁹.

Regardless of the functions of SH2 domains and similar protein modules, the initial event in tyrosine-kinase signalling must always be the phosphorylation of a tyrosine residue, be it autophosphorylation or the phosphorylation of an exogenous substrate. So one might expect PTKs to be tailored to phosphorylate physiologically relevant sites. This is the issue that Songyang *et al.* have addressed.

They have employed a degenerate peptide library, with the sequence shown in Fig. 1 but with a central tyrosine instead of serine, as a substrate for a variety of receptor and intracellular PTKs, and have used the procedure detailed above to investigate PTK substrate specificity. PTKs preferentially phosphorylate synthetic peptides with acidic residues amino-terminal to the tyrosine¹⁰, and indeed the PTKs tested all showed a preference for glutamic or aspartic acid at the -4 to -2 positions. While the receptors also preferred glutamic acid at the -1 position, the intracellular kinases preferentially phosphorylated tyrosine residues preceded by an isoleucine or valine. Most striking, however, is the selectivity shown by PTKs for residues carboxy-terminal to the phosphorylated tyrosine. All the PTKs tested showed a selection for hydrophobic residues at the +3 position. But although hydrophobic residues tend to dominate at all four positions carboxy-terminal to the tyrosine in peptides phosphorylated by receptors, Fps/Fes and Src kinases showed a preference for acidic residues in the +1 and +2 positions. How well do these predictions match reality, and what are the implications in physiological terms?

First, receptor autophosphorylation sites tend to be followed by hydrophobic residues, which serve as binding sites for the SH2 domains of signalling proteins such as phospholipase C- γ 1, phosphatidylinositol (PI) 3'-kinase, Ras GTPase activating protein, SH2-containing phosphatases, and so forth (so-called group III SH2 domains)⁶. This fits quite well with the specificity exhibited by receptors in the peptide library assay. The most striking example involves the insulin-receptor, whose principal substrate, IRS-1, has multiple YMXM motifs which, upon phosphorylation, bind the SH2 domains of PI 3'-kinase¹¹. In the degenerate library, the insulin-receptor

selects peptides with the sequence EEEYMMMM. To a first approximation, then, receptors appear to have a preference for phosphorylating sites that will bind group III SH2 domains.

The SH2 domains of intracellular PTKs generally recognize phosphotyrosine followed in the +1 and +2 positions by an acidic residue, with a hydrophobic residue at +3 (a feature of group I SH2 domains)⁶. Remarkably, the Fps/Fes, Src and Abl PTKs preferentially phosphorylate peptides that are well suited to bind their own

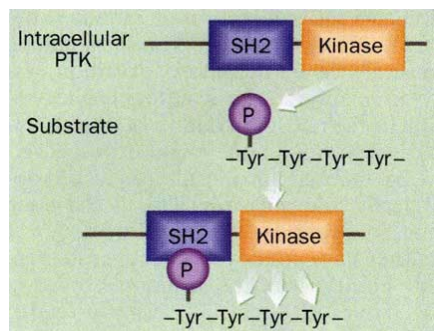


FIG. 2 Complementary functions of the kinase and SH2 domains of intracellular PTKs. The catalytic domains of kinases such as Src, Abl and Fps/Fes phosphorylate tyrosine motifs that subsequently bind with high affinity to the adjacent SH2 domain. In some cases, this may allow for repeated phosphorylation of the same molecule at several sites, as proposed for p130^{cas} phosphorylated by Abl^{1,14}.

SH2 domains. These results indicate that the SH2 domains of intracellular kinases may be involved in substrate recognition, as originally suggested, but in the sense that they bind proteins after phosphorylation by the neighbouring kinase domain. This might protect these substrates from dephosphorylation, hold them at a specific subcellular location, or allow for repeated phosphorylation of the same substrate, once it is tethered to the SH2 domain (Fig. 2). A good candidate for such processive phosphorylation is the p130^{cas} protein¹², which was originally identified by its ability to bind the SH2 domains of the Abl and Src kinases, and through its association with the Crk adaptor protein¹³. This protein has multiple Ile/Val-Tyr-X-X-Pro motifs, which in principle would be preferred substrates for the Abl kinase, and once phosphorylated should bind the Abl SH2 domain.

This raises the possibility that, following phosphorylation by the Abl kinase, p130^{cas} is captured by the Abl SH2 domain and is held in place for subsequent rounds of phosphorylation. There is evidence for this model, in that strong p130^{cas} phosphorylation by the Abl kinase seems to be dependent on the Abl SH2 domain¹⁴; in addition, the Crk adaptor protein, which binds to a proline-rich site in the Abl tail through its SH3 domain, contributes to processive p130^{cas} phosphorylation by

supplying, in *trans*, another SH2 domain to which p130^{cas} can bind. Like Abl SH2, the Crk SH2 domain favours pTyr-X-X-Pro sites. This suggests a novel role for SH2/SH3 adaptor proteins in modifying the specificity of intracellular PTKs.

Songyang *et al.* have an additional piece of data to support their cause. Homology modelling suggests that the +1 residue of a peptide substrate fits into a pocket lined by a methionine in most receptor PTKs, possibly accounting for their preference for hydrophobic amino acids at this position. Cytoplasmic PTKs, which select acidic residues at this position, have a threonine in place of the methionine. A substitution of the methionine in the presumptive +1-binding pocket of the Ret PTK for a threonine is associated with the human familial cancer multiple endocrine neoplasia type 2B (ref. 15), suggesting that this cancer might result from an alteration in Ret substrate specificity to resemble that of cytoplasmic PTKs^{1,16}.

These results are useful for thinking about signalling by PTKs. In particular, the substrate specificity of PTK kinase domains matches rather well with the binding specificity of SH2 domains, indicating that tyrosine kinases and SH2 domains have co-evolved, the former to phosphorylate sites that will be bound by the latter. The techniques pioneered by Songyang *et al.* will probably lead to a tremendous expansion in our knowledge of the substrate specificity of protein kinases. With increasing numbers of protein-kinase structures being solved, their recognition of substrates can be set on a structural footing. Similarly, as genomic and complementary DNA sequences become available, it becomes a simple job to search the database for likely protein kinase substrates. Ultimately, it may be possible to predict the substrate specificity of a protein kinase, and identify potential substrates, entirely on the basis of sequence information. □

Tony Pawson is at the Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada.

1. Songyang, Z. *et al.* *Nature* **373**, 536–539 (1995).
2. Crews, C. M., Alessandrini, A. & Erikson, R. L. *Science* **258**, 478–480 (1992).
3. Kemp, B. E., Pearson, R. B. & House, C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7471–7475 (1983).
4. Marais, R., Wynne, J. & Treisman, R. *Cell* **73**, 381–393 (1993).
5. Songyang, Z. *et al.* *Curr. Biol.* **4**, 973–982 (1994).
6. Songyang, Z. *et al.* *Cell* **72**, 767–778 (1993).
7. Pawson, T. & Schlessinger, J. *Curr. Biol.* **3**, 434–442 (1993).
8. Sadowski, I., Stone, J. C. & Pawson, T. *Molec. cell. Biol.* **6**, 4396–4408 (1986).
9. Mayer, B. J. & Baltimore, D. *Molec. cell. Biol.* **14**, 2883–2894 (1994).
10. Hunter, T. *J. biol. Chem.* **257**, 4843–4848 (1982).
11. Sun, X. J. *et al.* *Nature* **352**, 73–77 (1991).
12. Sakai, R. *et al.* *EMBO J.* **13**, 3748–3756 (1994).
13. Mayer, B. J. *et al.* *Nature* **332**, 272–275 (1988).
14. Mayer, B. J. *et al.* *Curr. Biol.* (in the press).
15. Hofstra, R. M. W. *et al.* *Nature* **367**, 375–376 (1994).
16. Santoro, M. *et al.* *Science* **267**, 381–383 (1995).

DAEDALUS

Stifled sound

THESE days we are surrounded by unwanted sound. Traffic roars on every road, Muzak pullulates in every public space, burglar and car alarms wail neglected in ever more urban areas, and personal radios and tape-players are everywhere. So Daedalus is inventing a super sound absorber, using the well-known aerogel process. This swells a solid to a tenuous gel in liquid carbon dioxide, then takes the system above its critical point and bleeds off the CO₂ as gas. The resulting expanded solid can be almost as light as air. DREADCO's chemists are now turning rubbers and elastomers into light and tenuous aerogels. Their low density gives them an acoustic impedance well matched to air, while their high elastic damping absorbs any sound traversing them.

DREADCO's Silentgel will be ideal for earplugs, cavity-wall fillings, acoustic tiles and sonic absorbers of all kinds. But Daedalus also sees the need for a selective sound absorber, to cut out maddening background noises but let through the loud direct sounds we want or need to hear. For this purpose, he is marketing Silentgel in the subtle form of a 100-nm coating on a harder backing. This thickness matches the typical vibrational displacement of a soft sound. Impinging on the thin coating, such a sound loses almost all its energy in compressing and dilating the flaccid film. By contrast a loud sound, with a displacement of 1 µm or more, is hardly absorbed at all. It flattens the coating completely, or dilates it fully, with 10 per cent of its cycle or less. The rest is reflected or transmitted without loss.

Sonic Subtractor film will clarify our sound-sabotaged lives. It will cut out the general sonic sludge, but let the useful information through. In the form of earplug diaphragms, it will present you with clear and distinct sounds only: close speech intended for your ears, or music that you have actually chosen to listen to. Similarly, in wallpaper form it will wonderfully sharpen a lecture theatre or restaurant. It will reflect the main speech, but will absorb the distracting background hubbub.

Subtractor film will have a special effect on sound of intermediate loudness. It will damp the low-amplitude parts of its cycle, but pass its peaks unattenuated. The resulting 'crossover distortion' is sometimes produced by cheap amplifiers. It is extraordinarily annoying. So a background sound which slowly gains intensity will be very obvious by its sudden grating harshness. A waiter surreptitiously turning up the Muzak will be immediately detected. David Jones

Origins of photosynthesis

SIR — Photosynthesis is of great antiquity; an essentially modern carbon cycle has existed at least since 3.5 ± 10^9 yr ago, and perhaps before 3.8 ± 10^9 yr ago¹. As a process, photosynthesis is extremely complex, and it is difficult to envisage how natural selection could have produced the intermediate stages in its evolution. We propose a model based on the premise that pre-photosynthetic life was associated with hydrothermal systems, in which certain early organisms developed photo-

taxis using near-infrared radiation, and that photosynthesis arose subsequently from bacteriochlorophyll-like molecules used for phototaxis.

The concept that the earliest eubacteria were thermophilic has considerable molecular support². Many archaea are thermophiles or acidophiles, well suited to hydrothermal conditions, so that it is not unreasonable to envisage an early hydrothermal world populated by a diverse community of bacteria^{3,4}.

Hydrothermal systems would have been widespread on the early Archaean Earth, both in shallow water around arc and plume volcanoes and also on mid-ocean ridges⁴, and could have supported widespread communities of chemotrophic life consisting of various archaea as well as primitive thermophilic eubacteria.

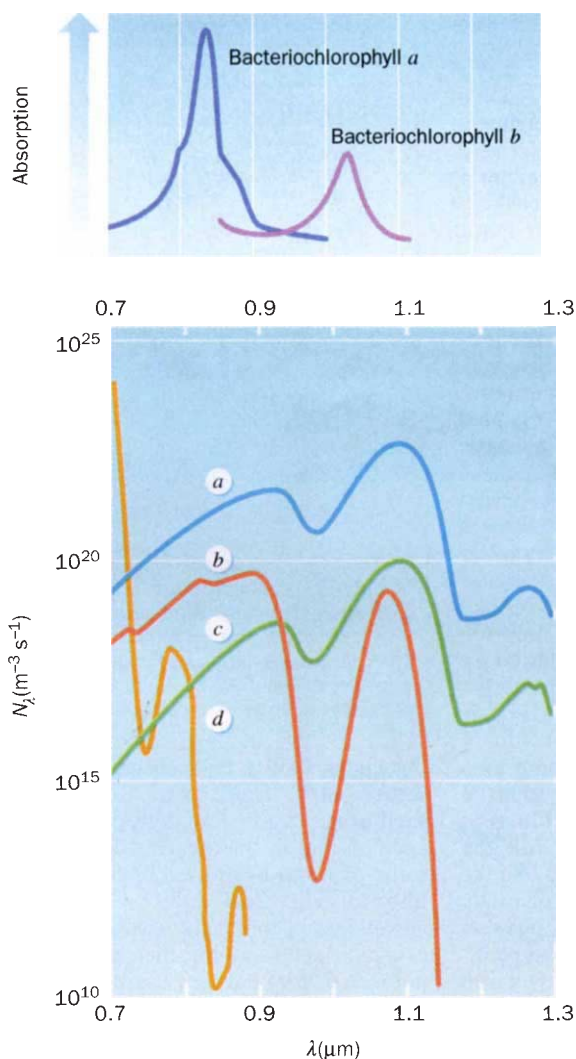
Critical to all organisms depending on hydrothermal systems is finding the optimum position for life in such complex environments, where outflow temperatures can range from a benign few degrees C to a hostile 400 °C, and fluid chemistry can show a great range of toxicity. This problem is made more difficult by the fluid flow surrounding the more active vents, where cold sea water moves inwards radially until it encounters the edge of the hot, turbulent buoyant plume of toxic fluid. Horizontal temperature gradients on the edge of individual plumes locally can reach hundreds of degrees in a few centimetres, and yet such places can provide optimal access to the chemical energy that supports life. Organisms that lose touch with such vents risk starvation, whereas those that come too close risk being cooked or poisoned. Modern vent organisms probably use a variety of sensors to control their position: one shrimp, *Rimicaris exoculata*, has highly derived eyes that may be modified as photosensors to detect light emitted from the hot vent water⁵. Phototaxis is

known in the purple bacterium *Rhodospirillum rubrum*⁶, which moves towards infrared light and away from visible light.

The spectrum of thermal radiation given off by a hot body depends on its temperature. As the temperature rises the intensity of radiation increases, and its maximum moves to shorter wavelengths. A body at 400 °C is just visible to a dark-adapted human eye (ref. 7, and J. R. C., unpublished observations). Eyes depend for their vision on rhodopsin, with a maximum sensitivity at 550 nm. Bodies at 400 °C are much brighter in the near infrared than at wavelengths appropriate for rhodopsin. This increased brightness is offset in water by increased absorption of radiation by water at wavelengths longer than the visible. Intense absorption does not set in, however, until about 950 nm, and there is a window of lower absorption at about 1,100 nm. The figure shows the calculated spectrum resulting from the interplay of the thermal emission from hot plumes and absorption by different thicknesses of water: there are two bands, 800–950 and 1,000–1,150 nm, where detection of thermal radiation would be optimal, and where sunlight filtered by a few metres of water is much less intense than the thermal emission. Light made up of thermal emission, and perhaps other components, has been detected at hydrothermal vents in the oceans⁸.

Intriguingly, the absorption spectra of bacteriochlorophyll match these bands closely. Bacteriochlorophyll *a* absorbs strongly in the 800–950-nm band, whereas bacteriochlorophyll *b* absorbs in addition in the region 1,000–1,150 nm (see figure). An early eubacterium capable of movement, able to detect near infrared in either band, would be able preferentially to occupy an optimum habitat in water more than a few metres deep. Bacteriochlorophyll may be even older than chlorophyll^{9,10}. It may be that an early chemotrophic bacterium possessed a primitive precursor of bacteriochlorophyll that absorbed light around 780 nm, allowing rudimentary detection of hot plumes. The selective advantage of thermotaxis could have driven the development of the use of the optimal 800–950-nm and 1,000–1,150-nm bands. Rudimentary photosynthesis using infrared from the hot vents could have begun in this setting, as a supplement to chemotrophy¹¹. We envisage evolution of an early precursor of bacteriochlorophyll in a chemotrophic organism for phototaxis.

Once evolved, a phototactic bacterium, possessing bacteriochlorophyll, which happened to be in a shallow water hydrothermal system, could make use of the infrared part of the spectrum of sunlight. It could begin facultative photosynthesis using H₂ and H₂S provided by the ambient hydrothermal setting. In shallow water, ultraviolet would be a problem, and



Top, Absorption spectra of bacteriochlorophylls *a* and *b*, measured in living bacteria, from ref 12. Bottom, Calculated spectra, shown as specific photon flux (N_λ , photons $\text{m}^{-3} \text{s}^{-1}$) against wavelength in μm . Spectra *a*, *b* and *c* are calculated for black-body emission from a 10-cm-diameter cylindrical source as modified by absorption in water (the absorption spectrum for water is taken from the Smithsonian meteorological tables). Spectrum *a* (blue) is for a source at 650 K (377 °C) viewed at 15 cm from the centre of the plume, spectrum *b* (red) is for a source at 650 K (377 °C) viewed at 50 cm from the centre of the plume; spectrum *c* (green) is for a source at 500 K (227 °C) viewed at 15 cm from the centre of the plume; and spectrum *d* (orange) is calculated for sunlight viewed at a depth of 10 m below the water surface (solar irradiance spectrum from ref. 13).

other pigments may have evolved as protection, later to be used with chlorophyll in developing the array of carotenoids and phycobiliproteins used in cyanobacterial photosynthesis. Simultaneously, along the different evolutionary line of the archaea, pigments in halobacteria were used as proton pumps, ancestral to rhodopsins in eukaryotes such as in modern vertebrate eyes.

We see photosynthesis as being an accidental by-product of thermal detection in an organism that was originally chemotrophic. Unlike the 'photosynthesis-first' hypothesis, our hydrothermal hypothesis invokes smaller evolutionary steps driven by natural selection. In an organism that is already surviving by chemotrophy, a smaller evolutionary change to allow infrared phototaxis gives local advantage. This then preadapts the organism for the next step: first for optional long-wavelength photosynthesis based on bacteriochlorophyll; and later for the development of chlorophyll to allow the use of visible light to split water.

Euan G. Nisbet

Department of Geology,
Royal Holloway College,
University of London,
Egham TW20 0EX, UK

Johnson R. Cann

Department of Earth Sciences,
University of Leeds,
Leeds LS2 9JT, UK

Cindy Lee Van Dover

Duke Marine Laboratory,
Pivers Island,
Beaufort,
North Carolina 28516, USA

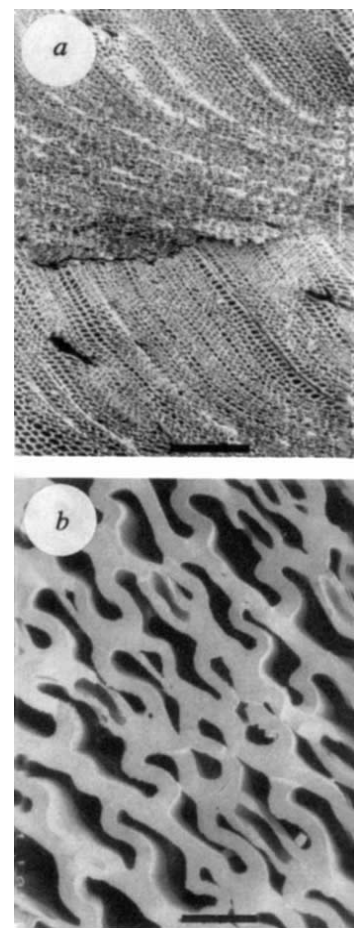
First use of coal

SIR — Charcoal analysis¹⁻⁴ from two Palaeolithic settlements (Les Canalettes⁵, Les Usclades⁶, south Massif Central, France) shows deformations never observed before in an archaeological context. Our observation of these deformations by scanning electron microscopy (*a* and *b* in the figure) suggested various possibilities to us. Fresh or dry wood could have been crushed either before or after being charred. We have investigated these hypotheses by an experimental approach, in which we attempted to restore similar deformation patterns to those observed in the remains.

We were successful only in the case of hot, dry and wet wood — compression of charcoal failed due to the brittle nature of the cell walls. Consequently, the crushed patterns can be explained only either as the result of highly technical or accidental action on wood material, or as having been induced by slow loading during very long periods (simulated by high temperature in the laboratory)⁷⁻⁹. The first explanation is incompatible with the proportion of crushed specimens observed. We therefore suggested an alternative hypothesis, that the wood was degraded geologically, before its use by prehistoric man.

Coal is essentially fossil peat, buried and subjected to increasing temperatures and pressure compressing its anatomical structure¹⁰. The anatomy of less altered charcoal shows that the wood looks like the genus *Pinus*, although we cannot tell if it is a pre-Pleistocene genus (photo *a*). From the anatomical observations, the remains could be fusains¹¹. The Jurassic coal outcrops within a radius of 7 to 15 km from the settlements in our study area are a possible source for the charred material (see ref. 5). The observation of Jurassic coal shows similarity with the most compressed remains (*a* and *b* in the figure). We still have not found the origin of the less compressed remains, but several outcrops have not yet been sampled.

Coal has been used for manufactured



a, Electron microscope photograph of archaeological charcoal with irregular compression in the transverse direction, mainly in early wood. The initial arrangement of the cell wall is strongly compressed but without cell wall fracture. *b*, Enlarged view of photo *a* (compressed cells of late wood).

objects in several Palaeolithic European settlements¹². There is no evidence for such usage in Les Canalettes. The presence of coal in one hearth suggests that some of it has been used as fuel by prehistoric man. Coal does present many advantages in comparison with wood, but how did the Palaeolithic men know of this potential? Why did they use it instead of wood? It could be that they discovered coal accidentally when wood supply was scarce during the last glaciation. The environment of Les Canalettes during the last glacial (73,500 ± 6,000 years before present) was dominated by large open landscapes on the plateau and there were wood resources in the valley. We observed more coal residues in the layer corresponding to the coolest period and less-developed arborescent vegetation. Moreover, some coal found in a neighbouring settlement, Les Usclades, enables us to correlate the disappearance of coal in the last occupation layer (layer 4: 8,220 ± 70 years before present) with forest development at the beginning of

- Schidlowski, M. *J. Southeast Asian Earth Sci.* **5**, 333-337 (1991).
- Achenbach-Richter, L. et al. *System. appl. Microbiol.* **10**, 231-240 (1988).
- Nisbet, E. G. *Nature* **321**, 206 (1986).
- Nisbet, E. G. *The Young Earth* (Allen & Unwin, London, 1988).
- Van Dover, C. L., Szuts, E. Z., Chamberlain, S. C. & Cann, J. R. *Nature* **337**, 458-460 (1989).
- Ragatz, L., Jiang, Z.-Y., Bauer, C. & Gest, H. *Nature* **370**, 104 (1994).
- Pelli, D. G. & Chamberlain, S. C. *Nature* **337**, 460-461 (1989).
- Van Dover, C. L., Delaney, J. R., Smith, M. & Cann, J. R. *EOS* (abstr.) **69**, 1498 (1988).
- Burke, D. H., Hearst, J. E. & Sidow, A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7134-7138 (1993).
- Blankenship, R. E. *Photosynth. Res.* **33**, 91-111 (1992).
- Van Dover, C. L. et al. *EOS* **75**, 44-45 (1994).
- Pfennig, N. in *Autotrophic Bacteria* (eds Schlegel, H. G. & Bowien, B.) 97-116 (Springer, Heidelberg, 1989).
- Weast, R. C. (ed.) *Handbook of Chemistry and Physics* 53rd edn (CRC, Cleveland, Ohio, 1972-73).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references.

- Vernet, J.-L. *Paléobiologie Continentale* **4**, No. 1 (1973).
- Thiébaud, S. *Documents d'Archéologie Française* **15**, 7-110 (1983).
- Heinz, Ch. *Paléobiologie Continentale* **16**, 1-212 (1990).
- Chabal, L. *L'Anthracologie, de l'Echantillonnage des Charbons de Bois à l'Interprétation du Paysage* (in the press).
- Meignen, L. *Monographie du CRA*, No. 10 CNRS, Paris, (1993).
- Maury, J., Frayssenge, H. & Renier, J. M. *Cahiers d'Archéologie Aveyronnaise* **5**, 13-23 (1991).
- Kollman, F. F. P. & Coté, W. A. *Wood, Science and Technology* Vol. 1 (1968).
- USDA Forest Service *Wood Handbook: Wood as an Engineering Material* (1987).
- Inoue, M. *J. Jap. Wood Res. Soc.* **39**, 867-875 (1993).
- Colin, R.W. (ed.) *Coal Geology and Coal Technology* (Blackwell Scientific, Oxford, 1983).
- Jones, T.P. & Chaloner W.G. *Palaeogeogr. Palaeoclim. Palaeoecol.* **97**, 39-55 (1991).
- Shepherd, R. *Prehistoric Mining and Allied Industries* (Academic, London, 1980).

the present interglacial. This suggests that prehistoric man may have resorted to coal at specific times of climate and forest regression, while using wood for preference.

I. Théry*, J. Gril†, J. L. Vernet*, L. Meignen‡, J. Maury§

*Laboratoire de Paléobotanique, Environnement et Archéologie, URA CNRS 1477,

†Laboratoire de Mécanique et Génie Civil, URA CNRS 1214, Université de Montpellier II, 34095 Montpellier Cedex 05

‡ERA 28, Préhistoire et Technologie, CRA-CNRS Sophia-Antipolis 06560 §13 Cité Beausoleil, 12740 Sébazac, France

A tombstone in Alzheimer's?

SIR — Harrington and Colaco in *News and Views*¹ considered recent intriguing studies providing evidence for glycation of amyloid- β protein (A β) and tau protein in Alzheimer's disease. A β and tau form the insoluble filamentous aggregates that comprise plaques and neurofibrillary tangles, respectively, in the brains of Alzheimer patients. Harrington and Colaco suggested that glycation of A β and tau may be causally involved in the neurodegeneration associated with A β and tau.

Nevertheless, additional findings to those discussed by Harrington and Colaco suggest that glycation of A β probably occurs after crosslinking of the peptide and after A β has damaged neurons. First, synthetic A β forms fibrillar aggregates in the absence of sugars (without glycation)^{2,3}. Second, synthetic (unglycated) A β can damage and kill neurons in culture, and its neurotoxic potency is correlated with its ability 'spontaneously' to form aggregates^{2,4}. Third, when synthetic A β is placed in physiological saline (in the absence of sugars) it generates free radical peptides that can directly damage enzymes⁵, and induce accumulation of reactive oxygen species in cultured

neurons^{6,7} and synaptosomes⁸. Fourth, studies of different A β fragments indicate that generation of free radical peptides correlates directly with aggregation kinetics (ref. 5). Fifth, although recent glycation connection studies^{9–11} provide strong evidence for the presence of advanced glycation end products in 'mature' plaques and tangles (comprised of aggregated A β and tau, respectively), they did not reveal whether advanced glycation end-products are present in 'immature' plaques and tangles before formation of A β and tau aggregates. If glycation is causally involved in the neuronal injury and death process, then glycation would be expected to occur early in the disease process.

We propose that glycation of A β is a late event in the evolution of Alzheimer's disease and results from free radical generation by A β itself. In this view, age-related changes in the brain, including reduced energy availability and an oxidizing environment, combine to promote radicalization of A β . A β radicals promote covalent crosslinking of the peptide, and propagate radicals to cell membranes. The peptide-associated radicals would also promote crosslinking of the peptide to sugars, resulting in glycation.

Mark P. Mattson*, John W. Carney†, D. Allan Butterfield‡

*Sanders-Brown Research Center on Aging and Department of Anatomy and Neurobiology,

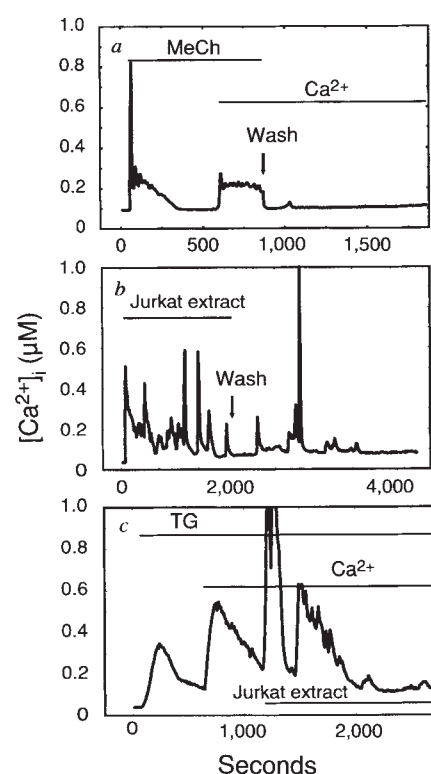
†Department of Pharmacology,

‡Center of Membrane Sciences and Department of Chemistry, University of Kentucky, Lexington, Kentucky 40536, USA

Calcium entry signal?

SIR — Randriamampita and Tsien¹ reported that acid extracts of Jurkat T cells contain a Ca²⁺-entry-stimulating factor that they concluded is the messenger for capacitative Ca²⁺ entry, and which they termed CIF (Ca²⁺-influx factor). Because of the potential significance of this discovery^{2–5}, we have further examined the effects of this material on Ca²⁺ signalling. Our findings demonstrate that this activity is unlikely to be the elusive cellular messenger for capacitative Ca²⁺ entry.

We prepared an extract of Jurkat cells following the procedure outlined in ref. 1, and tested its effects on fura-2-loaded 1321N1 astrocytoma cells. This extract caused an elevation of intracellular Ca²⁺ which was due to Ca²⁺ entry, usually accompanied by irregular oscillations (*b* in the figure). In unstimulated Jurkat cells, all the activity was in the particulate fraction of the cells, and on stimulation with phytohaemagglutinin, activity was



Calcium responses of 1321N1 astrocytoma cells to methacholine, Jurkat extract and thapsigargin. *a* and *b*, Single fura-2-loaded astrocytoma cells stimulated with either 100 μ M methacholine (initially in the absence of external Ca²⁺, with Ca²⁺ added to the medium where indicated); *a* or Jurkat extract (*b*). After establishing a sustained Ca²⁺ response in both instances, stimuli were removed by exchanging the bathing solution. Shown are representative traces from 3 similar experiments. *c*, After establishing a [Ca²⁺]_i response to thapsigargin (2 μ M) followed by addition of extracellular Ca²⁺ in a single fura-2-loaded astrocytoma cell, the Jurkat extract was added in the continued presence of thapsigargin. Shown are representative traces from 8 similar experiments. Ca²⁺ signalling in single astrocytoma cells was measured as described⁷. Jurkat extract was prepared as in ref. 1.

released into the soluble fraction. These findings reproduce those of Randriamampita and Tsien, and indicate that our Jurkat extract contains the same Ca²⁺-mobilizing principle described in their paper as CIF.

As Ca²⁺ signalling is a rapidly reversible response, so too should be the actions of intracellular mediators of Ca²⁺ signalling. But after washout of the Jurkat extract from astrocytoma cells, the [Ca²⁺]_i signals persisted for up to 30 min (*b* in the figure). By contrast, after washout of

- Harrington, C. R. & Colaco, C. A. L. S. *Nature* **370**, 247–248 (1994).
- Pike, C. J. et al. *J. Neurosci.* **13**, 1676–1687 (1993).
- Mattson, M. P., Tomaselli K. & Rydel, R. E. *Brain Res.* **621**, 35–49 (1993).
- Yankner, B. A., Duffy, L. K. & Kirschner, D. A. *Science* **250**, 279–282 (1990).
- Hensley, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3270–3274 (1994).
- Goodman, Y. & Mattson, M. P. *Expl. Neurol.* **128**, 1–12 (1994).
- Behl, C., Davis, J. B., Lesley, R. & Schubert, D. *Cell* **77**, 817–827 (1994).
- Butterfield, D. A. et al. *Biochem. biophys. Res. Commun.* **200**, 710–715 (1994).
- Vitek, M. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4766–4770 (1994).
- Smith, M. A. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5710–5714 (1994).
- Yan, S.-D. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7787–7791 (1994).

- Randriamampita, C. & Tsien, R. Y. *Nature* **364**, 809–814 (1993).
- Clapham, D. *Nature* **364**, 763–764 (1993).
- Putney, J. W. Jr & Bird, G. S. *J. Cell* **75**, 199–201 (1993).
- Putney, J. W. Jr *Science* **262**, 676–678 (1993).
- Meldolesi, J. *Curr. Biol.* **3**, 910–912 (1993).
- Kwan, C. Y. et al. *Am. J. Physiol.* **258**, C1006–C1015 (1990).
- Bird, G. S. J. et al. *Nature* **352**, 162–165 (1991).

methacholine (acetyl- β -methylcholine) the Ca^{2+} influx returned to basal levels within 1–2 minutes (a in the figure).

In agreement with ref. 1, we found that the extract-induced $[\text{Ca}^{2+}]_i$ signals tend to oscillate irregularly, while methacholine- or thapsigargin-induced signals do not. One explanation for this behaviour is that when the extract is applied the stores are full, and thus there may be mechanisms inactivating the Ca^{2+} -entry signal. But when we examined the response of astrocytoma cells to the extract following depletion of their intracellular stores with thapsigargin, we still saw large $[\text{Ca}^{2+}]_i$ oscillations (c in the figure on page 481).

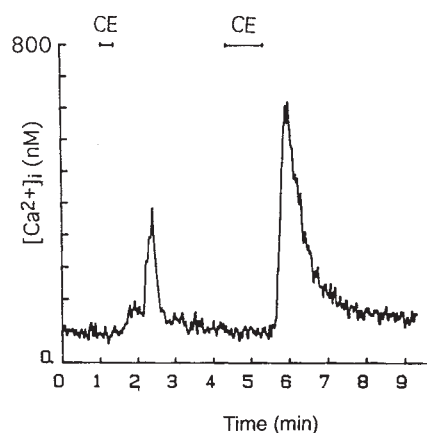
In the mouse lacrimal acinar cell, a cell type known to exhibit capacitative calcium entry⁶, the extract induced a clear intracellular mobilization of Ca^{2+} , as well as a sustained entry of Ca^{2+} (not shown). The response was entirely blocked by the intracellular application of the inositol trisphosphate receptor antagonist heparin, or by the external application of the muscarinic receptor antagonist atropine. The muscarinic-like activity is always found in the soluble fraction of Jurkat cells, indicating that this is a different principle from the one eliciting the Ca^{2+} entry responses in astrocytoma cells. When this muscarinic agonist activity was blocked, lacrimal cells did not respond to the Ca^{2+} -entry-stimulating principle in the Jurkat cell extract. However, when an extract of lacrimal cells is prepared by the same method as used for Jurkat cells, this extract induces a response in astrocytoma cells similar to the one induced by the Jurkat extract. Thus, although lacrimal cells contain the material, they do not respond to the activity in Jurkat extracts which was designated CIF.

Our findings indicate that the Ca^{2+} -signalling principle (or principles) in acidic extracts of Jurkat cells is unlikely to be the messenger for capacitative Ca^{2+} entry. It is not clear from our data whether the activity described by Randriamampita and Tsien¹ has a novel physiological role in Ca^{2+} signalling in certain cell types, or whether the extraction procedure has created a material which induces Ca^{2+} fluxes by non-physiological means.

**G. St J. Bird, Xiaopeng Bian,
J.W. Putney Jr**

Calcium Regulation Section,
Laboratory of Cellular
and Molecular Pharmacology,
Institute of Environmental Health Sciences,
NIH, National Institutes of Health
Research, Triangle Park,

RANDRIAMPITA AND TSIENT REPLY — We agree with Bird *et al.* that a crude cell extract is likely to contain several biological activities, not just the desired one. The muscarinic activity towards lacrimal cells found by Bird *et al.* is an example of just



Cytosolic free calcium ($[\text{Ca}^{2+}]_i$) in a single fura-2-loaded 1321N1 astrocytoma cell in a normal medium with 1 mM Ca^{2+} . During the two periods labelled CE, the cell was superfused with Jurkat cell extract with zero added Ca^{2+} and 1 mM EGTA, delivered from a local micropipette.

such contamination. Lacrimal cells may also be relatively insensitive to externally applied CIF, because cell types do vary in their sensitivity¹. However, the activity of Jurkat extract on the astrocytoma cells that we use for our assays is not due to this muscarinic component, because our unpublished results show that the $[\text{Ca}^{2+}]_i$ elevations are unaffected by atropine doses that completely inhibit muscarinic activation. We take the evidence for multiple activities as an incentive to purify the key messenger rather than to abandon the entire concept. Indeed, we have three independent chromatographic methods for fractionating cell extract and have seen resolution of multiple activities. The most active fraction has at least 20-fold more activity per unit dry mass than crude cell extract. A conclusive demonstration of the CIF concept will require purification to homogeneity and structural elucidation. The activity is not an artefact of acid extraction, because it is also releasable by sonication, but it then undergoes degradation in a few minutes², presumably by enzymes. Okadaic acid inhibits such degradation, which may explain why okadaic acid can potentiate Ca^{2+} elevations^{2,3}.

The figure (above) shows that the action of Jurkat extract can be briefer and less oscillatory than the results shown by Bird *et al.* During the periods labelled CE, single astrocytoma cells were superfused with cell extract containing zero Ca^{2+} . Because CIF action requires extracellular Ca^{2+} (ref. 1), $[\text{Ca}^{2+}]_i$ did not increase until the flow was stopped, which re-exposed the cells to normal medium with 1 mM Ca^{2+} but without cell extract. Then a sizeable $[\text{Ca}^{2+}]_i$ elevation appeared with negligible spiking or oscillation and quickly decayed. Our interpretation is that some CIF entered the cells during the super-

fusion, displayed its activation of Ca^{2+} influx once external Ca^{2+} was available, then decayed by leakage or metabolism, its source having been removed. In support of an intracellular site of action, Thomas and Hanley have shown that Jurkat extract works when microinjected into *Xenopus* oocytes⁴. Although many aspects remain mysterious, we think it premature to conclude that cell extract lacks an activity mediating capacitative Ca^{2+} entry.

Clotilde Randriamampita

Laboratoire de Neurobiologie,
École Normale Supérieure,
75005 Paris,
France

Roger Y. Tsien

Department of Pharmacology
and Howard Hughes Medical Institute,
University of California,
San Diego, La Jolla,
California 92093-0647, USA

1. Randriamampita, C. & Tsien, R.Y. *Nature* **364**, 809–814 (1993).
2. Randriamampita, C. & Tsien, R.Y. *J. Biol. Chem.* **270**, 29–32 (1995).
3. Parekh, A.B., Terlau, H. & Stühmer, W. *Nature* **364**, 814–818 (1993).
4. Thomas, D. & Hanley, M.R. *J. Biol. Chem.* (in the press).

Science and religion

SIR — The report “A Line in the Sea”, by J. A. Yoder *et al.* (*Nature* **371**, 689; 1994) not only attempts to explain the open ocean front phenomenon, but it also sheds light on a religious issue concerning a revealed natural phenomenon.

The Holy Book of Islam, al-Qur’an, has explained in two verses (XXV: 53 and LV: 19), the natural phenomenon called the “two adjacent seas” (Maraj al-Bahrain). According to the verses in the Qur’an, these two seas exist side by side with a boundary preventing them from mixing. Indeed, it specifies that the composition of water in these seas, especially the salt concentrations, are quite different.

For years, Islamic scholars have had difficulty in interpreting the concept of Maraj al Bahrain. Although diverse interpretations have been offered by different scholars and a variety of natural phenomena have been used to exemplify this revealed concept, none has been satisfactory from a scholarly perspective. It seems that the advancement of knowledge and technology and recent scientific achievements have made it possible for a better understanding of this natural phenomenon brought to the attention of man through the Qur’an.

Ebrahim Yazdi

21 Touraj Lane,
Valiasr Avenue,
Tehran 19666, Iran

Cruising for peace

Ziauddin Sardar

Rivers of Eden: The Struggle for Water and the Quest for Peace in the Middle East.
By Daniel Hillel. Oxford University Press: 1994. Pp. 355. \$30, £18.99.

THE region surrounding the 'twin rivers' of Tigris and Euphrates, and the mighty rivers of the Nile and the Indus, is the birthplace of civilization. It was here that the principal grain and leguminous crops were first bred and where tools still basic to agriculture in much of the world today were developed. From here came the first flint sickle, set in horn or bone, the olive mortar and the wine press, the baker's oven, then the plough, the wheel, and the pulley to draw water. Here our ancestors first domesticated plants and animals, tilled and irrigated fields and orchards, established villages, towns and cities, developed architecture and built monuments, and organized communities that transformed themselves into empires. Here ceramics and metallurgy were invented and writing and mathematics discovered. And here lies the seat of all the great religions of the world that have shaped the moral and ethical ideals of humanity. For centuries, this region was the world's main source of food. During the zenith of Muslim civilization, the curve that contains Egypt, Syria and Iraq was known as the 'fertile crescent'; and in Roman times the fertile crescent provided the Mediterranean civilizations with the bulk of their food and was described as the 'breadbasket of Europe'.

How has a region so rich in water and greenery, that once led the world in innovation and civilization, been transformed into a cultural and agricultural wasteland? Why has the legendary site of the Garden of Eden become so degraded and its people so embittered? Daniel Hillel seems to have been wrestling with these questions all his life. As an expert in plant and soil sciences, he has spent most of his academic and professional career studying the environmental problems of the Middle East. Now almost at the end of his long and distinguished career, he has found some insights.

The intuition came a few years ago when he joined a class of Bedouin children in Sinai. The children were being taught *adab*, manners or the "proper ways of the world" by a *mualem*, a traditional teacher. The *mualem*, who had no formal education, began by asking the children: "how much is one and one?". "Two", replied one of the pupils. "Sometimes it is so, sometimes not", replied the *mualem*, who then explained: "if you put one she-goat in a pen and then another, surely you will have two. But if you put one he-goat with another, they will fight and one might

get killed so that one and one would end up being one; or both might die so that one and one would be none. On the other hand, if you put one he-goat with one she-goat, you might eventually have three, or four or more. So, how much is one and one? That depends on circumstances and on inclination."

A modern scientist comes into contact with a life-enhancing tradition and is

when he takes the reader on a flight from Khartoum to Karachi and provides a running commentary on the devastation he sees below.

The arid and harsh sterility of the Sudan belies the fact that it is a land with a prodigious amount of water. The Sudd area of Southern Sudan contains the world's largest freshwater swamp, covering an expanse larger than the Everglades of Florida and the Okefenokee of Georgia. The salinated and waterlogged soil of the Nile valley contrast sharply with the fact that before it was dammed, the Nile's "annual pulsation and the associated fluctuation of the water-table created an automatically self-leaching cycle in which the salts were flushed away by the Nile itself". The Aswan Dam stops the Nile dead: its

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

"The Garden of Eden" (1659) by Jan van Kessel — now an agricultural wasteland.

humbled. Buried in the wise dictum of the old Bedouin master, Hillel found a truth of deep ecological implication. In the one-dimensional universe of Euclid's simplistic axiom, the whole is equal to the sum of its parts: one plus one equals two. But this holds true only when the parts are inanimate, noninteractive, sterilized. In real life, the whole includes not merely the sum of discrete parts but also the entire range of their complex interactions, both complementary and antagonistic.

The Middle East has been rendered sterile and life drained out of its soil by the application of one-dimensional science, a science that has excluded life, complexity and ecological and cultural interactions from its equations. If one needs water in this scheme, one simply uses a modern pump to drill deep tube-wells and quickly drain dry an entire aquifer dry. Or one dams a river, unconcerned about the havoc this causes to soil and social structures. Just what grandiose schemes, based on linear logic, have done to the Middle East is illustrated by Hillel

water has nowhere to go and so ends up underground, forcing the water table higher and higher and leading to waterlogged valleys and salination. Waterlogging is also causing the pyramids to sink and explains why bilharzia, caught by contact with a snail, has reached epidemic proportions in the region. Egypt is suffering an 'eleventh plague' to which it seemed immune for centuries.

Occasionally, however, there is a sight to behold. Over southern Iran, Hillel sees ancient 'chain wells' that have been used for irrigation purposes since Persian times. Known as *karez* in Persian and *qanat* in Arabic, these ingenious systems consist of one or more mother wells drained through a network of tunnels. A string of mounds surrounding the vertical shafts that provide access to the horizontal tunnels marks the surface path of the underground circuitry. For centuries before the arrival of the tubewells, the ecologically sound and the exceptionally durable *qanats* served as the principal means for supplying water for irrigation

and to villages and towns.

Hillel's main thesis is that the environmental degradation of the Middle East, and the resulting literal evaporation of its water, is one of the main causes of its malaise and conflicts. Water is not just the lifeblood of the region; it is also an integral part of the region's worldview and culture. Through an enchanting history of the region's civilization, and shrewd biblical and Qur'anic scholarship, *Rivers of Eden* highlights the crucial role of water in the development of the Middle East as well as in shaping its outlook. Without water, it seems, none of the 'hydraulic civilizations' would have emerged and the cultures of the Tigris-Euphrates, the Nile, and the Indus rivers would have perished long ago.

Just how deeply water is etched on Middle Eastern cultures can be judged by the fact that the Egyptians named their country *khami* ('black') after the dark loamy soil of the Nile Delta, which differed distinctively from the light-hued sands of the surrounding desert. As the fertile soil was considered to be the source of all life, the Greeks borrowed its name for the term *khemia*, which, after being processed by the Arabs, became 'chemistry'. The word for water in Persian, and the first word in the Persian dictionary, is *ab*. From *ab* comes *abad*, meaning abode, and *abadan*, meaning civilized. Civilized people make orchards called *paii-daeza* which leads to 'paradise'. Water features predominantly in the Old and New Testaments and is one of the main themes of the Qur'an where it is mentioned 60 times (along with 40 references to the sea and numerous indirect references in words representing fountains, springs, rain and clouds). It seems that no civilization, city or religion was conceivable without water.

The destruction of its environment and water-based resources has stripped the Middle East of its protective, insulating layer of tradition and civic society. Peace remains elusive in the region because the basic sources of the suffering of its people remain unrecognized and unaddressed. The root cause of conflict in the Middle East, Hillel argues, is the destruction of the region's traditional way of life through ill-fitting modernization and environmental degradation. Unable to sustain their traditional lifestyles and their ecological methods of farming, irrigation and use of water, millions of farmers and rural folk are displaced from their land. They gather in overcrowded cities without infrastructure, adequate housing, sanitation or employment. Deprivation and bitterness breed extremism.

Given today's trends, and the fashion for building dams, the situation could become worse. Turkey's Southeast Anatolia Project calls for the construction of 13 hydroelectric dams on the upper reaches

of the Tigris and Euphrates, including the massive Atatürk Dam on the Euphrates. When the project is completed, both Iraq and Syria will lose much of their current supply of water from the Euphrates: as Turkey extracts more water upstream, less remains for the two countries downstream. Syria would lose 40 per cent of its water supply from the Euphrates, Iraq would lose up to 80 per cent. The consequences would be devastating. A similar fate awaits the damming of the River Jordan.

The situation is made worse, Hillel argues, by the precarious nature of the Middle Eastern states, which were deliberately designed as unstable systems by the European powers. The Kurds, for example, were divided between three different countries and much of the Lebanon was consciously handed to the minority Christians.

The Middle East is an intricate web of interrelated problems, none of which can

be solved in isolation or by the "purely capitalistic notions of a competitive market economy". By their very nature, these problems demand the involvement and cooperation of all the people and the states of the region as well as a transdisciplinary approach.

Rivers of Eden is enchanting, enlightening and erudite. Its interdisciplinary analysis of hydraulics and soil science illuminates the hidden crevices of history and the darker corners of politics and international relations. Hillel shows clearly that the price of further conflict is a sure catastrophe while the dividends of peace are immeasurable, and in doing so he echoes the wisdom of an old Arabic proverb: "one whose hand is in water is not like one whose hand is in fire". □

Ziauddin Sardar is Visiting Professor of Science and Technology Policy, Middlesex University, Trent Park, Bramley Road, London N14 4XS, UK.

Beyond the limits of culture

Tom Kirkwood

How and Why We Age. By Leonard Hayflick. Ballantine: 1994. Pp. 377. \$24.

MANY a scientist has found immortality but none by Woody Allen's preferred method — not dying. Avogadro lives on through his number, Planck his constant, Ohm his law. Leonard Hayflick, happily very much alive today, will be known to future generations for discovering the

Hayflick limit, a phenomenon that seems partly responsible for the certainty that Woody Allen's wish is one day going to be thwarted.

The Hayflick limit, as many will know, describes the fact that normal human cells divide only a finite number of times (around 50 or 60) when grown in a culture flask or dish. In 1961, Hayflick, together with Paul Moorhead, published a landmark paper demonstrating this fact and suggesting that here was a manifestation of ageing at the cellular level.

To savour the full originality of Hayflick and Moorhead's work, we need to recall that for the previous 50 years it had been the firm belief that "cells inherently capable of multiplying will do so indefinitely if supplied with the right [conditions]". The quotation is from the rejection letter signed by the editor of the prestigious journal to which they first submitted their work.

It was Alexis Carrel, working mostly at the Rockefeller Institute (now Rockefeller University) in New York, who in 1912 created and energetically sustained the impression that cells grew indefinitely outside the organism from which they were obtained. Carrel's culture of chicken heart cells was maintained for 34 years, to the accompaniment of enthusiastic media coverage, until it was terminated in 1946, two years after his death. By then, human cancer cell

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Eternal youth — Peter Pan floats away in a picture by Alice B. Woodward (1907), inspired J. M. Barrie's play about a boy who refused to grow up.

New in paperback

The Language Instinct by Steven Pinker. HarperPerennial, \$14. "Here at last is a marvellously readable book about language, written by a real expert. Pinker tackles with wit and erudition the kinds of questions everyone asks about language. . . [to which] he brings not only an expertise in linguistics and psychology and a wide knowledge of biology, but also an ability to understand the ordinary person's linguistic hang-ups and to shake them loose with gentle ridicule. . . With its wealth of examples, its flawless typesetting, its wide-ranging bibliography and its irresistible good humour, Pinker's book is certain to increase its readers' respect for the amazing natural phenomenon that the author and his colleagues have made their life's study (Christopher Longuet-Higgins, *Nature* 368, 360; 1994).

Are We Alone? Implications of the Discovery of Extraterrestrial Life by Paul Davies. Penguin, £5.99. In his preface to this slender book, which is based on lectures given at the University of Milan in 1993, the author mentions that an estimated 170 books speculating about the existence of extraterrestrials were published before 1917. He now attempts "to rekindle this debate, and place it in a modern scientific context, by charting what aspects of contemporary science, and of our belief systems in general, are at stake".

The Astonishing Hypothesis: The Scientific Search for the Soul by Francis Crick. Simon and Schuster, £6.99. "An account of this dramatic enterprise by one of the greatest living biologists, written with clarity and charm. . . We are guided through the delightful intricacies of the psychology of vision; attention and memory; the structure of the brain; and the workings of its neurons. . . All this leads to potentially important speculations, the upshot being that consciousness remains mysterious. But Crick will continue to contemplate the future of brain science — the detailed and comprehensive study of neurons" (Richard L. Gregory, *Nature* 368, 359; 1994).

The Naturalist in Britain: A Social History by David Elliston Allen. Princeton University Press, \$16.95, £11.95. First published in 1976, this pioneering book has established itself as a key resource for historians of science and an excellent introduction to natural history for the general reader.

lines, such as the famous HeLa line established in 1952, were becoming known and these were, and still are, growing indefinitely.

Against this background, those unfortunate whose cultures of normal cells invariably died out, as they must have done, found themselves in a fix like the courtiers in Hans Andersen's classic tale of "The Emperor's New Clothes". Persuaded by the trickster tailors that failure to see the nonexistent cloth was proof of being unfit for office, the courtiers played along with the deception. Conventional wisdom held that cultures died only because they were not properly cared for. Who was going to trumpet their failures?

Well, Hayflick and Moorhead did. But they took care to design an experiment of elegant simplicity to prove their point. By growing mixed cultures of old and young cells they showed that the old cells died while the young thrived. The old and young cells were of different sexes and could be distinguished by their chromosomes. Conditions were the same, so the demise of the older cells was not through contamination or lack of care.

Other laboratories were quick to confirm the existence of the limit. It was soon shown that the number of cell doublings that a culture could grow through declined with the age of the cell donor. This supported Hayflick's notion of a link to the ageing process. Yet, despite more than 30 years of research, we still do not know exactly what causes the Hayflick limit, nor do we yet know the real nature of the link to the ageing process. Incidentally, it now seems likely that Carrel's culture was, wittingly or unwittingly, topped up with fresh cells.

Ageing is more than a matter of just running out of cells. The brain, for example, depends on neurons that early in our lives stop dividing but then continue to work perfectly well for decades, until they too age and die. Similarly, Hayflick's book ranges well beyond the confines of the limit, providing a comprehensive review from one of the pioneers of modern biogerontology. Written in clear, nontechnical language, it is an excellent introduction to the scientific and demographic literature on this multifaceted subject.

The first part of the book looks at the elusive problem of satisfactorily defining ageing. Hayflick compares species that age and those, such as lobsters, that he suggests do not. It can be hard to tell, especially if one has to wait decades to find out. A valuable point that Hayflick makes well is how the notion of ageing can vary according to the level at which the process is viewed. For example, bristlecone pines and giant sequoias may be very long-lived, but none of their cells lives as long as a human neuron.

Following a fact-filled section on the demography of ageing in the United States, Hayflick treats us to a thorough analysis of how the ageing process is played out in the human body. It is here that we find his account of the discovery of the limit. He moves on to review the results of the famous Baltimore Longitudinal Study of Aging, which catalogues how a group of dedicated volunteers have aged through time. The main message is that there is no general pattern of ageing applicable to all our organs: a person's rate of ageing may vary significantly from what might be predicted from the averages.

Theories of ageing receive broad coverage, from evolutionary theories through to specific molecular theories. There is a growing consensus that ageing happens through gradual accumulation of random changes or errors, perhaps because of evolved limitation in organisms' maintenance systems. Hayflick agrees, and suggests that on the basis of this reasoning we may conclude that the ageing process is malleable, that we can understand how it happens and that perhaps we can tamper with it.

Woody Allen would no doubt approve. But let us not lose sight of the underlying message of Hayflick's book: biogerontology today is a respected science, but there is a great deal more research that needs to be done. Non-sensational, popular treatment of this emotive subject is to be warmly welcomed. □

Tom Kirkwood is in the School of Biological Sciences and Department of Geriatric Medicine, University of Manchester, 3.239 Stopford Building, Oxford Road, Manchester M13 9PT, UK.

Research with a vengeance

David Edgerton

The Rocket and the Reich: Peenemünde and the Coming of the Ballistic Missile Era. By Michael Neufeld. Free Press: 1995. Pp. 368. \$21.95, £19.99.

THE nuclear-tipped ballistic missile, the key weapon of both the United States and the Soviet Union in the Cold War, owed much to German-speaking scientists and engineers of the Weimar era. In the few years between the Second and the Third Reich, Germany was a great centre of artistic, musical, scientific and technical innovation. Weimar Germany saw, for example, an enthusiasm for spaceflight, and experimentation with rockets on a scale not found anywhere else. But the Third Reich was itself a great centre of

innovation, and it was the German army, the Nazi party and the *Schutzstaffel* (SS) that gave the world the ballistic missile.

The study of the relationship between science, technology and society in Nazi Germany is not merely a parochial concern of German historians but is central to world history. This is not simply a reflection of the importance of German technology: one cannot study science and technology in the Third Reich without raising dangerous, difficult and important general questions about the relationship between knowledge and power, the moral responsibility of scientists and engineers, and the relationship between modernity and brutality. Michael Neufeld (of the National Air and Space Museum at the Smithsonian Institution in Washington DC) has written the first complete history in English of the story of the German liquid-fuelled rocket programme, which created the A-4 (better known as the Vengeance weapon 2, or V2), and he does not shy away from these issues.

The story itself is simple and chilling. The German army, which was dominated by gunners, was looking for alternatives to heavy artillery which was banned by the Versailles Treaty, and in 1929 started a small research programme with solid-fuel rockets. In December 1932, Baron Werner von Braun joined the project (his father was a member of the right-wing cabinet) and quickly became the technical director. By 1934 all public technical and military discussion of rockets was stopped by censorship and unofficial projects were suppressed. By the end of 1934 an A-2 was successfully launched (the A-1 had been a failure), and by early 1936 an A-4 ballistic missile was specified on the basis of extrapolation from the performance of the famous Paris Gun of 1918.

In the same year the German army began the construction of a research-and-development facility at Peenemünde on the Baltic coast. Before war broke out, and before the development was complete, the army had decided to build a massive factory at Peenemünde, but it was not until October 1942 that an A-4 was successfully launched. Production, first at Peenemünde and then, after an air raid in August 1943, largely at the 'Mittelwerk' near Nordhausen, yielded nearly 6,000 missiles; the equivalent, it has been estimated, of no fewer than 24,000 fighter aircraft. In September 1944 the first V2 was fired in anger; from then on, Mittelwerk was turning out around 20 missiles a day. By that stage of the war, the German

arms plants were highly dependent on slave labour, supplied by the growing SS empire; Auschwitz, we should remember, was a factory as well as an extermination camp. The V-2s were built by Russians, Poles, Belgians, Frenchmen and others under the control of the SS. At least 10,000 lives were taken in the process; this compares with the 5,000 people (mostly Belgians, and not Londoners) killed by V2s themselves. Very crudely, it took two

surely have been seen as monuments to technological lunacy.

Instead, in combination, the bomb and the rocket became the key technologies of the Cold War; and their wartime antecedents became the models for the development of new military technologies. After the war, some 100 Peenemünde engineers went to the United States, to work for the US army and later NASA (National Aeronautics and Space Administration). Von Braun's team designed the Jupiter ballistic missile, which put the first US satellite into orbit in 1958, and later the Saturn booster. Von Braun and his colleagues argued that their ambition had always been spaceflight, and that they were involved in the V2 programme only because that was the only way to the stars. Neufeld will have none of this. For most of his career, von Braun was a designer of ballistic missiles, a task he carried out with extraordinary energy, dedication and intelligence; talk of spaceflight was an extracurricular amusement. He also argues that von Braun was fortunate in being briefly arrested during the war because it helped him to present himself later as an apolitical engineer who fell foul of the Nazis: in fact his arrest was part of an SS campaign to take over the project. Von Braun, although never an active Nazi, was a Party member from 1937 and a member of the SS from 1940.

The V2 was an extraordinary technical achievement; in the 13 years from the start of the army project to the first successful flight, many intractable problems of engine design, guidance and supersonic aerodynamics had to be solved. The organizational and political work required to keep a project of this sort going in the difficult circumstances of war was immense. And the production and deployment of the missile were themselves considerable feats. All this Neufeld ably documents from the remaining archival material. But his important book is rightly dedicated to the enslaved workers who built the first ballistic missile with the intention that their sacrifice is never forgotten. □

David Edgerton is in the History of Science, Technology and Medicine Group, Imperial College, Sherfield Building, London SW7 2AZ, UK.

■ See *Nature* **372**, 511 (1994) for Tom Gehrels's review of *Wernher von Braun: Crusader for Space*, a biography by Ernest Stuhlinger and Frederick I. Ordway III.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Cold War clan — from left to right, Dornberger (commander at Peenemünde), Axter (scientist), von Braun (with fractured arm) and Lindenberg (scientist).

human lives to make a V2, and each killed one person.

The total cost of the research, development and production, Neufeld estimates, was around US\$500 million (\$5 billion at 1990s prices), about one quarter of the cost of the Manhattan Project. And yet the destructive power of all the V2s was the equivalent of a single British air raid. The terrible irony is that we should be grateful to von Braun, Dornberger, Hitler, Speer and many others for their efforts. It should not be thought that this kind of irrationalism was confined to the Germans. The \$2 billion spent on the Manhattan Project produced two bombs whose destructive effects were no greater than those of contemporary fire raids on Japanese cities. Market forces would never have produced the V2 or the bomb; far-sighted and all-knowing strategists would have stopped both projects very early on. Had preparation for war ended in 1945, the ballistic missile and the bomb would

Imperial War Museum

The structural basis of specific base-excision repair by uracil-DNA glycosylase

Renos Savva^{*}, Katherine McAuley-Hecht[†], Tom Brown[†] & Laurence Pearl[‡]

^{*} Structural Biochemistry Section, Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, UK

[†] Department of Chemistry, University of Edinburgh, King's Buildings, West Main Road, Edinburgh EH9 3JJ, UK

The 1.75-Å crystal structure of the uracil-DNA glycosylase from herpes simplex virus type-1 reveals a new fold, distantly related to dinucleotide-binding proteins. Complexes with a trioxynucleotide, and with uracil, define the DNA-binding site and allow a detailed understanding of the exquisitely specific recognition of uracil in DNA. The overall structure suggests binding models for elongated single- and double-stranded DNA substrates. Conserved residues close to the uracil-binding site suggest a catalytic mechanism for hydrolytic base excision.

URACIL-DNA glycosylases (EC 3.2.2.3) are a ubiquitous, highly conserved and extremely specific class of DNA repair enzymes¹⁻⁹. Their biological function is the specific removal from DNA of the normal RNA base uracil^{1,10,11}. Uracil-DNA glycosylases (UDGs) hydrolyse the *N*-glycosidic bond linking the base to the deoxyribose sugar, leaving an abasic site, which is subsequently removed by a 5'-acting AP-endonuclease (apurinic/aprimidinic) and a deoxyribosephosphodiesterase¹², and the resultant gap filled in and sealed by the action of DNA polymerase and DNA ligase (Fig. 1).

Uracil arises in DNA by spontaneous deamination of cytosine, which is greatly accelerated by environmental mutagens such as bisulphite ions and nitrous acid, and by misincorporation of deoxyuridine triphosphate during DNA synthesis¹³. Approximately 200 cytosine deamination events per day would be expected in a genome of 10¹⁰ base pairs, and deamination occurs 4,000 times faster in single-stranded DNA, making it a significant event in actively transcribed genes and replication forks¹⁴. Deamination of cytosine generates a pro-mutagenic U:G mismatch, which if not corrected, leads to a C→T transition mutation in the next round of DNA synthesis¹³. Misincorporation of dUTP during DNA synthesis is minimized by the action of deoxyuridine triphosphatase, yet over 10,000 deoxyuridine incorporation events are believed to occur per replicative cycle in a genome of 10¹⁰ base pairs¹⁴. Although uracil can base-pair adequately with adenine, and is not directly pro-mutagenic when in a U:A base pair, uracil in place of thymine in regulatory DNA sequences can substantially disrupt specific protein binding^{15,16}. Thus, uracil in DNA presents a serious threat to the informational integrity and regulated expression of the genome, unless it is actively removed.

Uracil-DNA glycosylases (UDGs) have been identified in a wide variety of prokaryotic and eukaryotic organisms, and in the pox and herpes families of viruses. UDGase sequences range between 199 and 359 amino acids in length, with a conserved carboxy-terminal component of 199–220 residues. Amino-terminal extensions found in eukaryotic and viral enzymes vary enormously in length and composition, and are probably involved in subcellular localization¹⁷. UDGases show exquisite specificity for efficient excision of uracil from DNA, with weak activity against synthetic uracil analogues such as 5-fluorouracil and 5-hydroxyuracil^{18,19}. They excise uracil from single- and double-stranded DNA, with excision from single-stranded DNA occurring most rapidly²⁰. No activity has been detected against any normal DNA base, nor against uracil in RNA²¹⁻²³.

The structural basis of the specificity of this fundamental DNA repair enzyme has not previously been understood. We have determined the structure of the UDGase from herpes simplex virus type-1 (HSV-1) by X-ray diffraction, at 1.75 Å resolution. The HSV-1 enzyme is 39% identical to that of the human, and 49% identical to the *Escherichia coli* enzyme, and thus provides a good model for the structure of the enzymes from either of these sources. Structures of complexes with a trinucleotide (5'-p-dTdTt-OH-3') and with free uracil base have also been determined, enabling us to locate the active site of the enzyme, and to determine how uracil is specifically admitted into, and other bases excluded from, this site. We can also identify several key residues, conserved in all UDGases, which are probably involved in the catalytic mechanism. Modelling from the observed structures explains the specificity for DNA

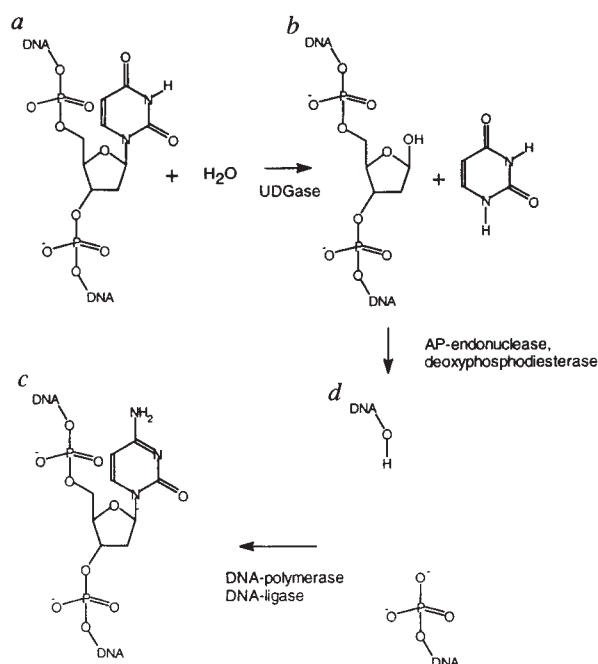


FIG. 1 Pathway of uracil-DNA base excision repair. Uracil base is removed by uracil-DNA glycosylase; the resulting abasic sugar is removed by apurinic/aprimidinic endonuclease and deoxyribosephosphodiesterase, and the resulting gap filled by DNA polymerase using a second strand as a template. Ligation by DNA ligase completes the repair process.

[‡] To whom correspondence should be addressed.

over RNA, and the overall structure of the enzyme suggests a binding and tracking mode for single- and double-stranded DNA substrates.

The enzyme from HSV-1 is of particular interest, as it is important for viral reactivation²⁴. HSV-1 spends extended periods of latency in mature human neurons that lack a UDGase activity. During these periods, which may be many years, the episomal viral genome accumulates large numbers of genetically damaging cytosine deamination events, which must be repaired by the UDGase during reactivation. Absence of this activity dramatically impairs viral reactivation and neuroinvasion²⁴. The structure of HSV-1 UDGase will aid the design of specific inhibitors that may be the precursors of new anti-herpes drugs.

Structure determination

The structure of HSV-1 UDGase was solved by multiple isomorphous replacement of a trigonal crystal form of the enzyme alone²⁵, and an orthorhombic crystal form containing a complex of the enzyme with a trinucleotide. Phase information from the two crystal forms was combined by real-space cross-map averaging and phase extension/combination, giving an electron density map for the DNA co-crystal, in which most of the polypeptide chain could be traced and the sequence unambiguously placed. This model was transformed to provide an initial model for the native structure. Structures for the native enzyme, and crystals of complexes with free uracil base and 5'-p-dTdT-OH-3' have been refined to high resolution. Crystallographic data are given in Table 1. The quality of the electron density can be judged from Figs 2c and 4c.

Overall architecture

HSV-1 UDGase consists essentially of a single domain, formed by a left-handed coil of four helices from residue 17 to 80, leading into an alternating β - α - β structure, from residue 81 to the C terminus at residue 244 (Fig. 2a). The first 16 amino acids show no electron density and are presumed to be disordered in the crystals. The second helix in the α -coil has prolines at residues 34 and 38, which give it a substantial curve, and a third proline at residue 43 causes a tight turn into the third helix. The fourth helix also has two dislocations in its hydrogen-bonding pattern caused by a proline at residue 77, and a tight β -conformation at residues 69 and 70. The twisted parallel sheet that lies at the heart of the subsequent β - α - β structure has a topology reminiscent of the symmetrical Rossman fold found in many enzymes involved in binding nucleotide coenzymes²⁶, but involves only four strands, and has a substantial elaboration consisting of a 25-residue meandering loop and two helices, between strands 1 and 2 (Fig. 2d). The overall structure resembles a slightly dented matchbox.

The distribution of charged groups on the protein surface is highly anisotropic, with a significant excess of basic residues on one of the large faces of the molecule. A sinuous channel runs across the surface of this face, with a distinct pocket towards one end, formed by the side chains of Tyr 90, Phe 101 and the main and side chains of Pro 86, Gln 87 and Asp 88. The amide side chain of Asn 147 forms the bottom of the pocket. All the residues involved in the formation of this pocket occur at the C-terminal ends of the first and second β -strands and are totally conserved in UDGase sequences (Fig. 3). Two clearly defined sulphate ions are bound in the channel on either side of the pocket.

Uracil base binding

UDGases are end-product-inhibited by uracil¹⁸. A difference map calculated from a native HSV-1 UDGase crystal soaked in uracil shows clear electron density corresponding to the structure of uracil, in the conserved pocket (Fig. 4a). The uracil placed in this electron density makes a number of hydrogen bonds with the side chain of the totally conserved Asn 147 and with the main-chain atoms of Gln 87, Asp 88 and Phe 101. The uracil

ring is stacked parallel to and in van der Waals contact with the plane of the side chain of Phe 101, and perpendicular to the phenyl ring of Tyr 90, with which the uracil C-5 and C-6 atoms are in van der Waals contact (Fig. 4b). The N-1 atom makes a hydrogen-bonding contact with a solvent molecule in the top of the pocket, defining the orientation of the ring. Binding of uracil causes a small movement of the side chain of Phe 101 towards the plane of the base, but otherwise there is very little change in the position of groups in the pocket on binding.

Thymine-trinucleotide binding

In the orthorhombic trinucleotide co-crystals, the DNA binds at one end of the channel with the 3' end directed towards the centre (Fig. 2b). The phosphate groups of the 3' and 5' nucleotides are in contact with the enzyme through hydrogen-bonding and/or charge-pair interactions. The sugar and base of the 5' nucleotide make no direct contact with the enzyme, and the base is involved in solvent-linked interactions with symmetry-related molecules in the crystal lattice. The middle thymine base packs against the hydrophobic side chains of Pro 213 and Leu 214, and its attached deoxyribose ring packs against the side chain of Pro 111. These highly conserved residues provide a sequence-independent binding site for this nucleotide. The 3' nucleotide is positioned well into the cleft, with the phosphate group interacting with the side-chain and main-chain atoms of the highly conserved Ser 112, at the N terminus of the α -helix from residues 112 to 124. This phosphate binds in the same place as a sulphate in the native and uracil co-crystals. The thymine base of this nucleotide is directed towards the uracil-binding pocket, but fails to penetrate it, and the base remains at the mouth of the pocket, packed against the side-chain ring of Tyr 90, with the 5-methyl group in contact with the top edge of the ring of Phe 101, and the amino nitrogen hydrogen-bonded to the peptide oxygen of Asp 88 (Fig. 4c). All three nucleotides have 2'-endo sugar conformation and *anti*-glycosidic torsion angles, although the overall backbone conformation is far removed from B-form DNA, and the bases are completely unstacked. Trinucleotide binding causes little conformational change in the regions interacting with the 5' and middle nucleotides, but binding of the 3' nucleotide is accompanied by a 180° rotation of the side chain of Asp 88 in the wall of the uracil-binding pocket.

Basis of uracil specificity

Uracil-DNA glycosylases must have an exquisite specificity for the excision of uracil, but not any other base, from DNA, but not from RNA. Even a low level of activity against the four naturally occurring DNA bases would generate huge numbers of abasic sites and cause significant fragmentation of the DNA, with dire biological consequences, whereas activity against uracil in RNA would totally destroy the template function of messenger RNA. Uracil base binds to UDGase in a totally conserved pocket and makes a set of interactions involving its Watson-Crick base-pairing groups which are highly specific for uracil. The normal DNA bases adenine and guanine would be excluded from the pocket by the side chain of Tyr 90, which would clash with the imidazole ring of the purine bases and would in any event make unfavourable polar contacts in the pocket. Cytosine, although small enough to enter, would make highly unfavourable repulsive interactions with the polar groups lining the pocket, whose configuration is complementary to uracil alone. Thymine, with identical base-pairing groups to uracil, could in principle make the correct set of hydrogen-bonding contacts in the pocket, but would be excluded by virtue of its 5-methyl group, which would clash with the ring of Tyr 90. A second mechanism for exclusion of thymine may also be in operation. In the trinucleotide complex, the 3' thymine poised to enter the uracil pocket, instead makes a favourable set of interactions at the mouth of the pocket, which provides hydrophobic contacts for the 5-methyl group and the face of the thymine ring, and hydrogen bonds with protein and water, for the Watson-Crick

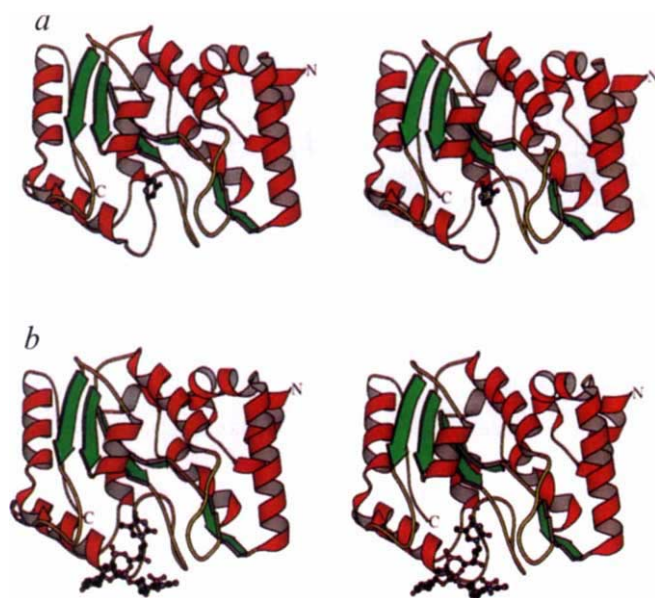


FIG. 2 Overall structure of UDGase. *a, b*, Stereo pairs showing the overall structure of UDGase. α -Helices are shown in red, β -strands in green and other structure in yellow. Figures were generated with MOLSCRIPT⁴⁰. The positions of free uracil base (*a*) and 5'-p-dTdT-OH-3' (*b*) in complex crystals are shown as ball-and-stick models. *c*, Electron density for the β - α - β segment from amino acids 182 to 208 in the native enzyme. The electron density is from a σ_A -weighted $2F_o - F_c$ map with calculated phases from the refined model, contoured at $1.0 \times \text{r.m.s. electron density}$. *d*, TOPS⁴¹ diagram of UDGase showing the topological arrangement of secondary structural elements. α -Helices are shown as circles, β -strands are shown as triangles. The strands forming the central parallel sheet are numbered in order of their position in the sequence.

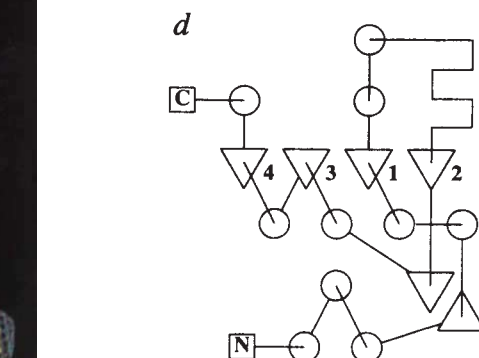
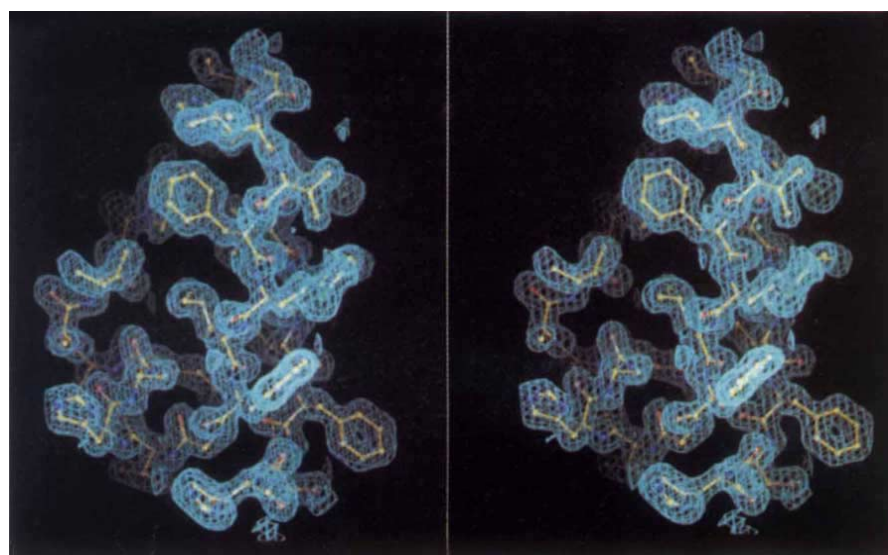


FIG. 3 Conservation and secondary structure in HSV-1 UDGase sequence. *a*, Stereo space-filling picture of UDGase looking towards the conserved pocket. Residues are coloured according to their conservation in homologous UDGase sequences, with most conserved as blue, varying at least conserved as red. Homology is based on multiple alignment of the UDGase sequences from herpes simplex virus type 1, Epstein-Barr virus, human cytomegalovirus, varicella zoster virus, human, *Saccharomyces cerevisiae*, *E. coli* and fowl pox virus (SwissProt Entries: UNG_HSV11, UNG_EBV, UNG_HCMVA, UNG_VZVD, UNG1_HUMAN, UNG_YEAST, UNG_ECOLI, UNG_FOWP1). *b*, Secondary structure and conservation of HSV-1 UDGase along the amino-acid sequence from the observable N terminus at Leu 17 to the C terminus. α -Helices are shown as spirals, β -strands as arrows. The 'bar' is coloured according to conservation of the residue in homologous sequences, as in *a*.

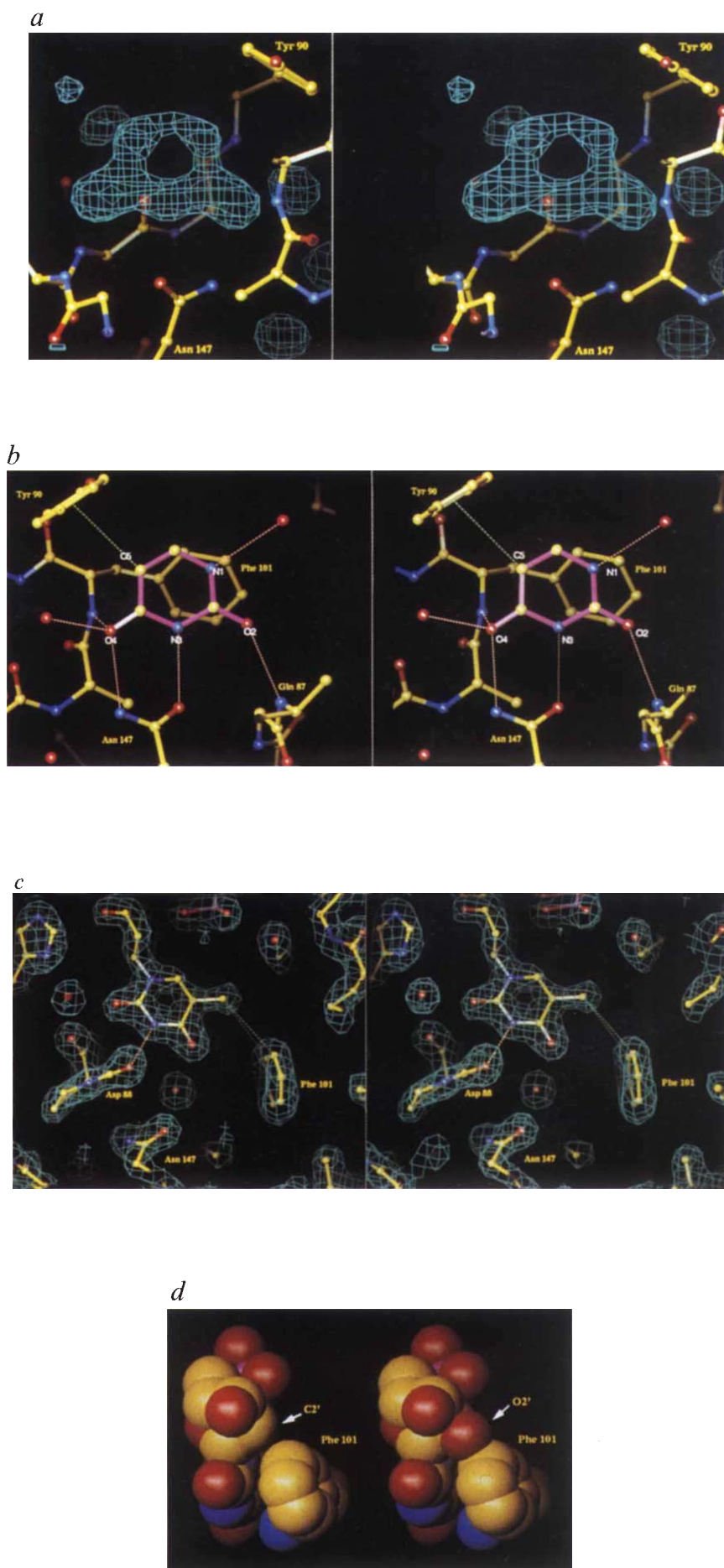
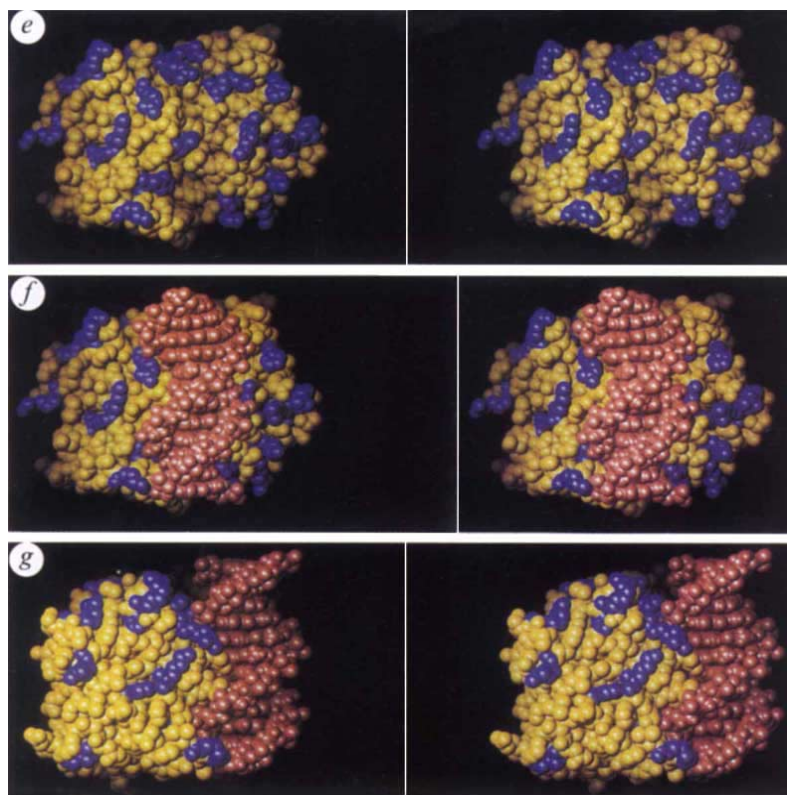


FIG. 4 Ligand binding. *a*, Stereo pair of electron density for uracil, in a σ_A -weighted $F_o - F_c$ map contoured at $3.0 \times$ r.m.s. density. *b*, Stereo pair showing interactions in the uracil-binding site; hydrogen bonds from the N-1, O-2, N-3 and O-4 atoms of uracil are shown as red/white dashed lines, and the close van der Waals contact between the C5 of uracil and the ring of Tyr 90 is shown as a green/white dashed line. Atoms are coloured by type: carbon, yellow; oxygen, red; and nitrogen, blue. The bonds of the ligand are highlighted in magenta. *c*, Stereo pair showing binding of the 3' thymidine in the trinucleotide complex. Hydrogen bond between the uracil N-1 and the carbonyl oxygen of Asp 88 is shown as a red/white dashed line, and the van der Waals contact between C-5 of uracil and the ring of Phe 101 is shown as a green/white dashed line. The electron density is from a σ_A -weighted $2F_o - F_c$ map with calculated phases from the refined model, contoured at $1.25 \times$ r.m.s. electron density. *d*, Space-filling model of productively bound 2'-deoxyribouridine (left) and ribouridine (right) based on observed coordinates for the phosphate and uracil base. The 2' hydroxyl of ribouridine (red) clashes sterically with the side chain of Phe 101. *e*, Space-filling model of HSV-1 UDCase with basic residues highlighted in blue. The view is the same as Fig. 2*a*, *b*. The DNA-binding channel runs from top to bottom of the molecule. *f*, Proposed model for double-stranded DNA binding. *g*, As *f*, but rotated around the vertical axis to show the concave face of the major groove of the DNA packing against the complementary convex surface of the protein.



groups on the base. This alternative binding site may act as a specific 'trap' for thymine, preventing even a low level of potentially productive binding in the pocket.

Specificity for DNA

Although in the present study we did not directly observe the binding of a uridine nucleotide, the positions of the productively bound uracil base, and of the 5'-phosphate group of the non-productively bound 3' thymidine, provide reliable reference points for accurate modelling of a productively bound uridine nucleotide. Only a 2'-*endo* sugar pucker and *anti*-glycosidic torsion angle satisfy the conformational constraints imposed by the observed positions of the base and phosphate. In this conformation, the 2' carbon of the deoxyribose sugar is in comfortable van der Waals contact with the edge of the side chain of Phe 101 (Fig. 4d). A ribose sugar in this position, with a hydroxyl group attached to the 2' carbon, would generate a steric clash with Phe 101, and prevent productive binding of the uracil base in the pocket. This interaction provides a satisfactory mechanism for the primary resistance of uracil-RNA to excision by UDCase, although interactions in peripheral sites may also contribute to a weaker overall interaction with RNA than with DNA.

A model for binding of extended DNA substrates

Both trinucleotide and free uracil bind at one end of the channel that runs across the predominantly basic face of the protein. A longer single-stranded substrate could run along the full length of the channel, making essentially sequence-nonspecific interactions, similar to those with the central base in the trinucleotide complex. UDCase are unusual in displaying base-excision activity against both single- and double-stranded uracil-DNA substrates¹⁶, and must be able to bind either. A double-stranded dodecamer of B-DNA can be docked surprisingly well against the basic face of the enzyme, with the sugar-phosphate backbone at the 5' end of one strand running through the end of the channel, close to the trinucleotide- and uracil-binding sites, while the other strand interacts with the channel at the opposite end.

The short α -helix formed by residues 190–196 protrudes slightly from the face of the enzyme and provides a complementary surface to the concave face of the major groove, at the centre of the DNA-dodecamer (Fig. 4g). The sugar-phosphate backbones of both DNA strands interact with many polar and positively charged residues across the surface of the enzyme. The degree of complementarity between the enzyme and the DNA duplex can be increased somewhat by gently bending the DNA towards the body of the enzyme. Non-specific interactions with an extended double- or single-stranded DNA substrate would allow the enzyme to track along the DNA, until a uracil base is encountered.

In the observed mode of uracil binding, specific recognition occurs by hydrogen-bonding between the protein and the Watson-Crick base-pairing groups of the base. Excision of uracil from double-stranded DNA must therefore involve disruption of base-pairing, and a significant distortion of the backbone conformation of double- or single-stranded uracil-DNA substrates, for the uracil base to penetrate the binding pocket. Distortion of the backbone, even in duplex DNA, need not be extensive, and may just involve the scissile uridine 'flipping out' of the body of the double helix and into the uracil-binding pocket, with relatively little perturbation of the rest of the DNA. Such a

process has been observed in the interaction of *HhaI*-methyltransferase with its cognate DNA sequence²⁷. The need to break hydrogen bonds and de-stack bases so as to excise uracil from double-stranded DNA explains the observation that excision from single-stranded DNA occurs more rapidly and, in contrast to excision from double-stranded DNA, is independent of flanking sequence effects¹⁶.

Catalytic mechanism of uracil excision

Uracil-DNA glycosylases excise uracil by hydrolytic attack on the *N*-glycosidic bond connecting the base to the C-1' atom of the deoxyribose sugar. Close to the site of the bound uracil lie several residues that are absolutely conserved in all UDCase sequences, and which are within regions of the sequences implicated in substrate binding by crosslinking studies²⁸. Two of these residues, Asp 88 and His 210, have been implicated in catalysis by site-directed mutagenesis of their homologues in the human enzyme²⁹, and are within 'striking' distance of the C-1' carbon of a modelled, productively bound uridine nucleotide (see above). Asp 88 would be chemically suitable as a nucleophile, but is too far away to interact directly with the C-1' carbon without a substantial conformational change in the protein backbone. In the uracil complex, however, a well ordered water molecule is hydrogen-bonded simultaneously to the peptide carbonyl and side-chain carboxyl of Asp 88, and would be in van der Waals contact with the C-1' carbon of the deoxyribose ring. This water molecule is ideally placed for attack on the C-1' carbon, and would be an excellent nucleophile owing to its interaction with the carboxyl of Asp 88, which could act as a general base and abstract a proton from the bound water. Non-productive binding of thymidine in the trinucleotide complex displaces this water molecule and causes the side chain of Asp 88 to rotate away from the substrate. This effectively 'disarms' the enzyme and may be a further factor in the rejection of thymidine as a substrate.

Nucleophilic attack by water is unlikely to be sufficient in itself to displace uracil, and some other factor would be required to improve the leaving group quality of the base. This might be

TABLE 1 Results of structural analysis

a, Crystallographic data

Crystal	Space group	λ_{collect} (Å)	Unique	Mult.	Complete %	d_{min} (Å)	R_{merge}
Native crystals							
UDG-1	$P3_1$	1.5418	8,039	3.7	99.6	2.5	0.077
UDG-2	$P3_1$	0.88	19,073	2.4	82.4	1.75	0.055
UDG 1 + 2 merge	$P3_1$	—	22,566	—	95.9	1.75	0.119*
Co-crystals							
UDG + dTdTdT-1	$P2_12_12_1$	1.5418	8,669	3.3	95.4	2.5	0.061
UDG + dTdTdT-2	$P2_12_12_1$	0.800	21,694	3.6	90.3	1.8	0.048
UDG + d ^U UdT	$P2_12_12_1$	1.5418	8,558	3.4	95.8	2.5	0.060
UDG + dT ^d UdT	$P2_12_12_1$	1.5418	10,259	5.2	80.1	2.2	0.050
UDG + d ^U d ^U dT	$P2_12_12_1$	1.5418	3,791	3.0	72.8	3.0	0.106
UDG + d ^B UdT	$P2_12_12_1$	1.5418	8,656	3.7	94.9	2.5	0.036
UDG + dT ^B UdT	$P2_12_12_1$	0.9199	6,868	5.6	60.9	2.3	0.061
Soaked crystals							
UDG + Uracil	$P3_1$	0.87	21,944	2.1	94.2	1.75	0.029
UDG + K ₃ IrCl ₆	$P3_1$	0.9199	7,571	2.8	96.0	2.5	0.040
UDG + K ₂ OsCl ₆	$P3_1$	1.5418	5,312	1.3	74.2	2.8	0.054
UDG + PCMBs	$P3_1$	1.5418	4,791	1.2	83.4	2.8	0.038

b, Phasing statistics

Derivative	N sites	d_{phas} (Å)	Δiso	R_{cullis}	Phasing power
Orthorhombic crystals					
d ^U d ^U dT	2	3.0	0.109	0.55, 0.60	1.1, 1.6
dT ^d UdT	1	2.5	0.088	0.68, 0.74	0.8, 1.2
d ^U dTdT	1	2.5	0.087	0.76, 0.76	0.8, 1.1
dT ^B UdT	1	2.5	0.062	0.86, 0.87	0.5, 0.7
d ^B UdTdT	1	2.5	0.060	0.81, 0.81	0.6, 1.0
<FOM> (8.0–2.5 Å) = 0.58					
Trigonal crystals					
K ₃ IrCl ₆	1	2.5	0.267	0.92	0.6
K ₂ OsCl ₆	1	2.5	0.161	0.98	0.3
PCMBs	1	2.8	0.122	0.95	0.4
<FOM> (8.0–2.5 Å) = 0.26					

c, Refinement statistics

Crystal	Resolution (Å)	No. of atoms	R	R_{free}	Δbond (Å)	Δangle (°)
Native	8.0–1.75	1,918	0.201	0.267	0.013	1.96
Native + uracil	8.0–1.75	1,925	0.196	0.239	0.014	2.00
Native + dTdTdT	8.0–1.8	2,035	0.172	0.219	0.015	1.93

Abbreviations: a , λ_{collect} , wavelength of X-rays used in data collection; Unique, number of symmetry unique reflections measured; Mult., number of observations/number of symmetry unique reflections; Complete, fraction of possible symmetry unique reflections actually measured in the resolution range; d_{min} , spacing of highest angle diffraction data; R_{merge} , $\sum \sum_i |I(h) - \langle I(h) \rangle| / \sum_i I(h)$ where $I(h)$ is the mean intensity after rejection of outliers; b , d_{phas} , Resolution limit of phasing; Δiso , mean isomorphous difference $\sum ||F_{\text{PH}}| - |F_P|| / |F_P|$; R_{cullis} , $\sum ||F_{\text{PH}} - F_P| - |F_H|| / |F_{\text{PH}} - F_P|$; Phasing power, $\sum |F_P| / \sum (|F_P| e^{i\theta} + F_H) - |F_{\text{PH}}|$, where $|F_{\text{PH}}|$ and $|F_P|$ are the observed derivative and native amplitudes, respectively, F_H is the calculated heavy-atom structure factor, and θ is the calculated phase angle. Summations are over all observations. Pairs of values of R_{cullis} and phasing power for orthorhombic crystals are for the centric and acentric reflections, respectively. c , R , Conventional crystallographic R factor calculated on 95% of all data in resolution range, used in refinement of models; R_{free} , R factor calculated for 5% of data omitted from refinement; Δbond and Δangle , r.m.s. deviations from ideal values for bond lengths and angles. **Crystallization, data collection and processing.** HSV-1 UDGase was expressed and purified as described²⁵ from a sequence corresponding to nucleotides 10,156 to 10,890 of the HSV-1 UL-2 open reading frame, cloned into the vector pTrc99a (Pharmacia-LKB Biotechnology) in *E. coli* JM109. The sequence encodes a fully active 244-amino-acid product containing the region homologous to other UDGases but lacking the highly variable N-terminal leader sequence. The structure was solved using data from two crystal forms; the previously reported trigonal crystals (space group $P3_1$ or $P3_2$, $a = 65.14$ Å, $c = 49.16$ Å) and co-crystals with a single-stranded triphosphonucleotide 5'-p-dTdTdT-OH-3' grown under essentially the same conditions (space group $P2_12_12_1$, $a = 46.65$ Å, $b = 61.83$ Å, $c = 86.67$ Å). Crystallization conditions were modified from those originally reported to: 54 mg ml⁻¹ protein, 5.2% PEG 8000, 34 mM Na₂SO₄ and 17 mM sodium phosphate buffer, pH 6.8. Uracil complex crystals were obtained by soaking native crystals in stabilizing buffer containing 20 mM uracil for 4 days. All data were recorded on MAR Research Image Plate detectors with derivative crystals carefully aligned for data collection to maximize the number of Friedel pairs recorded on the same image (except for the hexachlorosmate data which was recorded on an Enraf-Nonius FAST Area Detector, with no anomalous data). Diffraction images were processed using MOSFLM³¹ or MADNES³² and reduced and scaled using the ROTAVATA/AGROVATA, TRUNCATE and FHSICAL programs of the CCP4 suite³³. **Phasing.** Conventional soaking experiments with the native crystals gave three derivatives of low phasing power. The IrCl₆³⁻ (1.5 mM, 4 weeks) and OsCl₆³⁻ (1.2 mM, 5 days) anions both bound close to each other and in proximity to a sulphate-binding site in the native crystals, while the PCMBs (*p*-chloromercuribenzenesulphonate, slow increase 1→10 mM over 48 h and 12 h at 10 mM) bound to Cys 204. Derivatives of the trinucleotide co-crystals were produced by co-crystallization with a series of trinucleotides in which the 5' and central bases were substituted by 5-iodouracil and 5-bromouracil in various combinations. These derivatives were completely isomorphous with the original co-crystals and displayed significant phasing power, despite the small size of the heavy-atom contribution. Data for the IrCl₆³⁻ derivative and one of the 5-bromouracil substituted co-crystals were collected at a wavelength chosen to optimize the anomalous scattering signal. Heavy-atom positions were determined from isomorphous difference Patterson functions using the VECSUM program (I. J. Tickle) and common hands and origins were established from cross-phased difference Fourier. Heavy-atom parameters were refined and MIR phases calculated using the CCP4 implementation of MLPHARE³⁴, and modified by automated solvent flattening and histogram equalization using the DM program (K. Cowtan, unpublished data). Maps at 3.0 Å for both crystals showed clear solvent boundaries, and the map for the DNA co-crystals displayed some recognizable secondary structure features, but was substantially uninterpretable. Maps were significantly improved using cross-crystal averaging in the following way. First, the rotation relating the molecule in the DNA co-crystal to that in the native crystal was determined from the observed amplitudes with the CCP4 program ALMN, giving a clear solution at 9.6 × the function r.m.s. for data from 8 to 2.5 Å. Electron density for a single molecule in the 3.0-Å DNA-co-crystal map was isolated using a logical mask defined interactively with O (ref. 35), and transformed, by the rotation function solution, into the trigonal cell of the native crystal using RAVE³⁶. Normalized partial structure factors from the transformed electron density were calculated and used in a T2 translation function in both possible trigonal space groups, using the CCP4 TFCC program. A single solution at 9.8 × the function r.m.s. was obtained in space group $P3_1$, while the T2 function in $P3_2$ was featureless at 3.0 × the function r.m.s. The rotation and translation function solutions were refined by maximizing the electron-density correlation in real-space with RAVE, and the two maps were subjected to repeated cycles of averaging and phase combination, with the data progressively extended to 2.8 Å. Finally, calculated phases from the averaged oligonucleotide co-crystal map at 2.5 Å resolution were combined with the MIR phases and an electron-density map was calculated. This map was readily interpretable, and allowed building of more than two-thirds of the structure. The remainder of the protein structure and the bound trinucleotide were identified in σ_A weighted³⁷ $2F_o - F_c$, combined phase maps. Protein coordinates from the trinucleotide co-crystals were transformed by the previously determined rotation and translation function solutions, to give an initial model for the native structure. **Refinement.** Models for the native enzyme and the uracil and trinucleotide co-crystals were refined by simulated annealing refinement in X-PLOR³⁸. The variation from ideality of geometric parameters is within the ranges expected for protein structures at this resolution³⁹, and more than 90% of residues have main-chain torsion parameters within the preferred regions of the Ramachandran plot. Only one residue, Phe 101 (which is extremely clearly defined in the electron density, and totally conserved in all UDGase sequences) has a 'forbidden' conformation ($\phi = 69^\circ$, $\psi = -38^\circ$). This residue forms a key part of the substrate-binding pocket and suggests that this energetically unfavourable conformation is nonetheless conserved because of the biological function it confers. The peptide bond between residues Leu 62 and Pro 63 is *cis*. All model building was performed using O (ref. 35).

* R_{merge} for UDG 1 + 2 is based on 4,437 reflections in resolution range 3.7–2.4 Å common to both datasets.

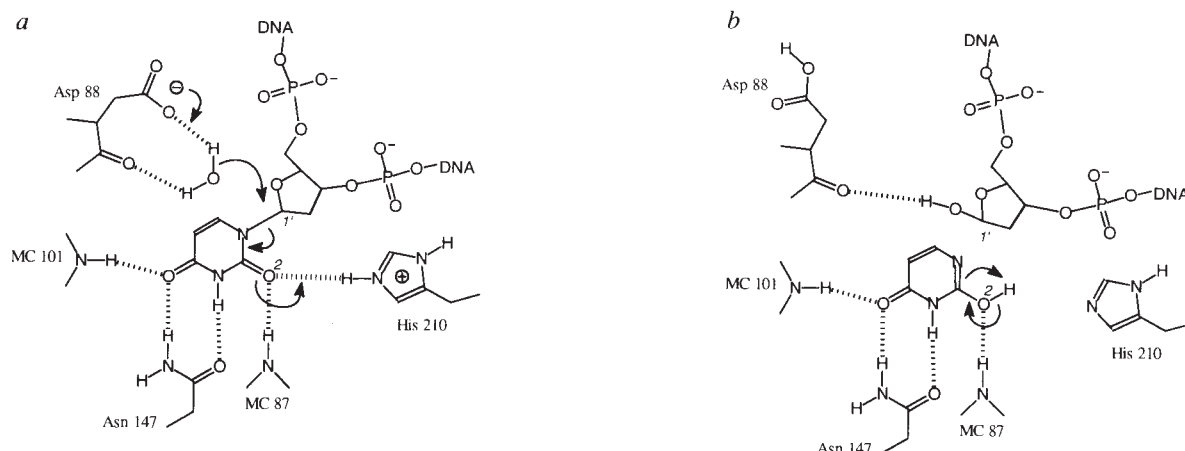


FIG. 5 Catalytic mechanism for hydrolytic base excision by UDGase. *a*, His 210 acting as a general acid protonates O-2 of the uracil base, shifting the tautomer from ketoamine to imineol, enhancing the leaving group qualities of the base. Concomitant attack on C-1' by a water mol-

ecule, activated by Asp 88 acting as a general base, leads to displacement of the uracil base. MC indicates main-chain atoms. *b*, The free uracil relaxes to the ketoamine tautomer. Release of products, reionization of the reactive groups, and binding of water, restores the enzyme.

achieved by distortion or, more probably, by protonation of the ring. Although direct protonation of N-1 would be difficult to achieve, as its lone pair is delocalized in the π -bonding of the ring, the same effect could be produced by protonation of O-2. The pyrimidine would then be much more readily displaced by a nucleophilic water. Unlike the other groups in the uracil ring, O-2 is relatively exposed and only involved in a single hydrogen bond to a peptide nitrogen. The other side of O-2 is directed towards the side chain of the totally conserved His 210, which would be an excellent source of a proton; however, the distance from the histidine ring nitrogen to O-2 is 5.1 Å, too large for a direct proton transfer. In oligonucleotides, uracil is only excised efficiently if there are at least two phosphate groups to the 3' end of the uridine nucleoside³⁰. Modelling from the observed positions of the trinucleotide and uracil base indicates that these phosphates and their intervening sugar and base would clash

sterically with residues in the loop carrying the conserved His 210, which must move to accommodate this part of the substrate. A movement of only 2 Å, triggered by binding of the 3' end of an extended substrate, would be sufficient to bring the side chain of His 210 close enough to protonate O-2 and effect hydrolysis (Fig. 5).

The structure of HSV-1 UDGase provides the first view of a highly specific monofunctional DNA glycosylase whose activity is essential for the maintenance of the genome in all DNA-containing organisms. The exquisite specificity of UDGase results from a combination of favourable pseudo-Watson-Crick interactions with uracil, and exclusion of the normal DNA bases. Exclusion of thymine, the DNA homologue of uracil, may also involve a specific 'trap'. Uracil is recognized by interactions with base-pairing groups, implying a 'flipping out' mechanism for excision of uracil in double-stranded DNA. □

Received 18 November 1994; accepted 9 January 1995.

- Lindahl, T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3649–3653 (1974).
- Kaboev, O. K., Luchinka, L. A. & Kuziakina, T. I. *J. Bact.* **164**, 421–424 (1985).
- Crosby, B., Prakash, L., Davis, H. & Hinkle, D. C. *Nucleic Acids Res.* **9**, 5797–5809 (1981).
- Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E. & Helland, D. E. *EMBO J.* **9**, 3121–3125 (1989).
- Upton, C., Stuart, D. T. & McFadden, G. *Proc. natn. Acad. Sci. U.S.A.* **90**, 4518–4522 (1993).
- Mullaney, J., Moss, H. W. McL. & McGeogh, D. J. *J. gen. Virol.* **70**, 449–454 (1989).
- Caradonna, S., Worrall, D. & Lirette, R. J. *J. Virol.* **61**, 3040–3047 (1987).
- Baer, R. et al. *Nature* **310**, 207–211 (1984).
- Davison, A. J. & Scott, J. E. *J. gen. Virol.* **67**, 1759–1816 (1986).
- Cone, R., Duncan, J., Hamilton, L. & Friedberg, E. C. *Biochemistry* **16**, 3194–3201 (1977).
- Lindahl, T., Ljungquist, S., Siegt, W., Nyberg, B. & Sperens, B. *J. biol. Chem.* **252**, 3286–3294 (1977).
- Dianov, G. et al. *Nucleic Acids Res.* **22**, 993–998 (1994).
- Lindahl, T. *Prog. Nucleic Acid Res. molec. Biol.* **22**, 135–192 (1979).
- Mosbaugh, D. W. *Rev. biochem. Tox.* **9**, 69–130 (1988).
- Verri, A., Mazzarello, P., Biamonti, G., Spadari, S. & Focher, F. *Nucleic Acids Res.* **18**, 5775–5780 (1990).
- Focher, F., Verri, A., Verzeletti, S., Mazzarello, P. & Spadari, S. *Chromosoma. Berl.* **102** (Suppl.), S67–S71 (1992).
- Slupphaug, G. et al. *Nucleic Acids Res.* **21**, 2579–2584 (1993).
- Mauro, D. J., De Reil, J. K., Tallarida, R. J. & Sirover, M. A. *Molec. Pharmacol.* **43**, 854–857 (1993).
- Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P. & Wallace, S. S. *J. biol. Chem.* **269**, 18814–18820 (1994).
- Eftedal, I., Guddal, P. H., Slupphaug, G., Volden, G. & Krokan, H. E. *Nucleic Acids Res.* **21**, 2095–2101 (1993).
- Cone, R., Duncan, J., Hamilton, L. & Friedberg, E. C. *Biochemistry* **16**, 3194–3201 (1977).
- Lindahl, T., Ljungquist, S., Siegt, W., Nyberg, B. & Sperens, B. *J. biol. Chem.* **252**, 3286–3294 (1977).
- Krokan, H. & Wittwer, C. U. *Nucleic Acids Res.* **9**, 2599–2613 (1981).
- Pyles, R. B. & Thompson, R. L. *J. Virol.* **68**, 4963–4972 (1994).

- Savva, R. & Pearl, L. H. *J. molec. Biol.* **234**, 910–912 (1993).
- Rossmann, M. G. et al. in *The Enzymes* Vol. 11 (ed. Boyer, P. D.) 61–102 (Academic, New York, 1975).
- Klimasaukas, S., Kumar, S., Roberts, R. J. & Cheng, X. *Cell* **76**, 357–369 (1994).
- Bennet, S. E., Jensen, O. N., Barofsky, D. F. & Mosbaugh, D. W. *J. biol. Chem.* **269**, 21870–21879 (1994).
- Krokan, H. E. et al. *Proc. N.Y. Acad. Sci. (Abstr.)* (July 31–Aug. 4, Burlington, VT, 1993).
- Varshney, U. & van de Sande, J. H. *Biochemistry* **30**, 4055–4061 (1991).
- Leslie, A. G. W. *MOSFLM Users Guide* (MRC-LMB, Cambridge, 1994).
- Pflugrath, J. W. & Messerschmidt, A. *Munich Area Detector New EEC System* (1992).
- Collaborative Computational Project No. 4. *Acta crystallogr.* **D50**, 760–763 (1994).
- Otwinowski, Z. in *Isomorphous Replacement and Anomalous Scattering: Proceedings of the CCP4 Study Weekend* 23–38 (SERC Daresbury Laboratory, 1991).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
- Kleywegt, G. J. & Jones, T. A. in *From First Map to Final Model: Proceedings of the CCP4 Study Weekend* (eds Bailey, S. et al.) 59–66 (SERC Daresbury Laboratory, 1994).
- Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
- Brunger, A. T. *X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
- Laskowski, R. J., MacArthur, M. W., Moss, D. S. & Thornton, J. M. *J. Appl. crystallogr.* **26**, 283–290 (1993).
- Kraulis, P. J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Flores, T., Moss, D. S. & Thornton, J. M. *Protein Engng* **7**, 31–37 (1994).

ACKNOWLEDGEMENTS. We thank our colleagues B. O'Hara, I. Sanderson, S. Wilson and C. Aguilar for assistance with data collection and helpful discussion; the Department of Crystallography at Birkbeck College and the Daresbury Laboratory for access to X-ray diffraction equipment; and J. Thornton, D. Jones, C. Orengo, A. Michie, P. Swann, I. Tickle and E. Dodson for discussion. L.P. particularly thanks F. Richardson-Pearl for discussion of enzyme mechanisms, and their son Michael for not being born until after this paper was submitted. This work was supported by the Cancer Research Campaign (L.P. and R.S.) and The Wellcome Trust (T.B. and K.M.H.). T.B. thanks the Royal Society of Edinburgh for a Caledonian Research Fellowship.

Inhibition of giant-planet formation by rapid gas depletion around young stars

B. Zuckerman*, T. Forveille† & J. H. Kastner‡

* Department of Physics and Astronomy, University of California Los Angeles, Los Angeles, California 90024, USA

† Groupe d'Astrophysique, Observatoire de Grenoble, 414 rue de la Piscine, F-38041, Grenoble Cedex, France

‡ Center for Space Research, Building 37-667a, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ALTHOUGH stars form from clouds of gas and dust, there are insignificant amounts of gas around ordinary (Sun-like) stars. This suggests that hydrogen and helium, the primary constituents of planets such as Jupiter and Saturn, are not easily retained in orbit as a star matures. The gas-giant planets in the Solar System must therefore have formed rapidly. Models of their formation generally suggest that a solid core formed in $\leq 10^6$ yr, followed by the accretion of the massive gaseous envelope in $\sim 10^7$ yr (refs 1–5). But how and when the gas of the solar nebula dissipated, and how this compares with the predicted timescale of gas-giant formation, remains unclear^{6,7}, in part because direct observations of circumstellar gas have been made only for stars either younger or older than the critical range of 10^6 – 10^7 yr (refs 8–15). Here we report observations of the molecular gas surrounding 20 stars whose ages are likely to be in this range. The gas dissipates rapidly; after a few million years the mass remaining is typically much less than the mass of Jupiter. Thus, if gas-giant planets are common in the Galaxy, they must form even more quickly than present models suggest.

A consonant hypothesis, that rocky terrestrial planets are common in our Galaxy, seems to be emerging among many astronomers and planetary scientists. The observational basis for this view includes discovery and study, by the IRAS satellite and by ground-based observers, of large quantities of particulate matter in orbit around many young stars and of a millisecond pulsar orbited by (presumably rocky) planets⁴². These particles represent the raw material for rocky and icy planets. Theoretical investigations often offer support for this hypothesis¹⁶. The commonality of terrestrial planets, should this indeed be the case, may be in part due to the ability of a star to store the appropriate raw materials (kilometre-sized planetesimals and larger objects) in orbit for very long periods of time. Thus the Earth may have taken $\sim 10^8$ yr to accumulate¹⁶.

On the other hand, neither theory nor observation paints a picture of solar system formation in which gas-giant planets

necessarily occur frequently. Wetherill⁴ has argued that the existence of Jupiter and Saturn may be essential for the existence of life on Earth because they cleansed the Solar System of most of its planetesimals (comets) that, otherwise, would be striking the Earth at a rate of the order of 1,000 times the actual impact rate. Perhaps gas giants are not very common; the reason that we see them in our Solar System is simply that, otherwise, we just would not be here.

To ascertain whether 'protoplanetary disks' contain raw material for the formation of Jupiter-like planets, we conducted a search for molecular gas towards a representative sample of young stars that IRAS found to be orbited by large quantities of dust (for example, refs 17–21). This dust can act as a shield, protecting orbiting molecules against photodissociation by the stellar and interstellar ultraviolet radiation fields¹⁵. Calculations²² suggest that CO and N₂ are by far the most abundant molecules other than H₂. Therefore almost all studies (for example, refs 8–15), including the present one, have involved the CO molecule (neither H₂ nor N₂ has permitted radio-frequency transitions). CS has also been studied²³.

Our particular choice of target stars was dictated by two primary considerations. First is proximity to Earth; because Solar-System-sized regions (radii < 100 AU) subtend only a small solid angle as seen from Earth, they are very difficult to detect at all with existing radio telescopes²⁴. Second is age; we want to investigate stars whose ages are in the range $(1\text{--}10) \times 10^6$ yr, presumably the time of giant-planet formation. Ages of young stars can be estimated with a variety of techniques that include; location in or near dark molecular clouds, placement along pre-main-sequence evolutionary tracks in a Hertzsprung–Russell diagram, atmospheric abundance of lithium, chromospheric activity, rotation and location of orbiting dust particles near a star.

In November 1993, January–February 1994 and October 1994 we observed 30 dusty stars at the frequencies of the $J=2\text{--}1$ and $1\text{--}0$ rotational transitions of CO with the 30-m radio telescope of the Institut de Radio-Astronomie Millimétrique (IRAM) on Pico Veleta, Spain. The 'front-ends' at both frequencies were SIS receivers and the spectrometers were 256-channel filter banks with resolutions of 1 MHz and 100 kHz. The beam sizes had been previously measured by IRAM staff to be 21 arcsec at 115 GHz and 11 arcsec at 230 GHz; the pointing uncertainty was ~ 3 arcsec. Backgrounds were removed by switching the sub-reflector at 0.5 Hz between a star position and a nearby reference position located, typically, 60 arcsec away. Additional data, obtained both by mapping around the position of the target stars and in frequency-switched mode, indicate that, even for HD32509, HD35187 and HD36112 (Table 1), which are located at the edge of the Taurus molecular clouds, interstellar CO does not confuse our spectra.

FIG. 1 Data for the dusty star TW Hydra, at the frequency of the $J=2\text{--}1$ rotational transition of CO. Abscissa, radial velocity with respect to the Sun. The spectral resolution is 100 kHz ($\equiv 0.13$ km s⁻¹). Ordinate, the brightness temperature averaged over the main beam of the 30-m telescope. If the rotation period of TW Hya is 1.3 days (ref. 40) and $V \sin i < 15$ km s⁻¹ (ref. 41 and B.Z., unpublished data), where V is the stellar rotational velocity and i is the inclination of the rotational axis, then the angle between the line of sight to the star and its rotation axis is $< 15^\circ$. This is consistent with the nearly face-on orientation of nearby gas implied by the narrow CO line.

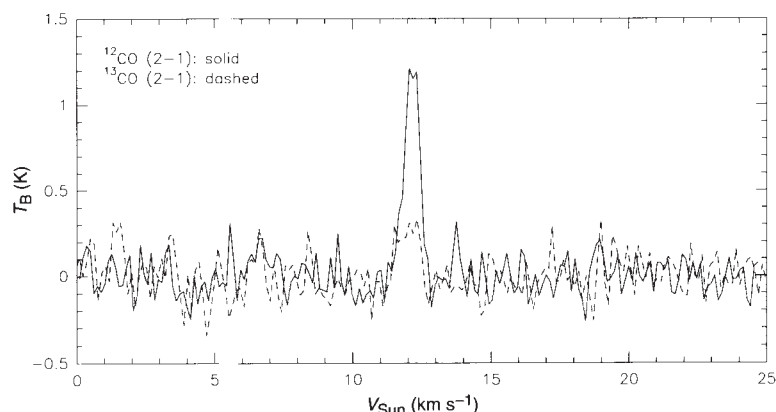


TABLE 1 Observed and derived quantities of young stars orbited by dust

Star	Spectral type	1950 RA (h min)	1950 dec. (°)	$V_{\odot}(\text{opt.})$ (km s ⁻¹)	$T_{\text{B}}(^{12}\text{CO})$ (mK)	$V_{\odot}(\text{CO})$ (km s ⁻¹)	ΔV (FWHP) (km s ⁻¹)	D (pc)	R (AU)	T_{CO} (at R) (K)	Assumed incl. (°)	Age (10 ⁶ yr)	M_{H_2} ($M_{\oplus}$)	M_{d} ($M_{\oplus}$)
49 Cet	A3V	1 32	-15	9	50	12	8	70	140	52	60	10	2	0.02-1
τ 1 Eri	F6V	2 42	-18	25	<50			14	80	40	60	>100	<0.06	0.01
HD32509	A2V	5 01	26		<50			150	150	50	60	6	<5	>0.2
HD34700	G0V	5 17	5		160	21	4.6	60	50	40	30	<10	80	1
HD35187	A2/3IV/V	5 20	24		≤ 120			150	150	50	60	5	<15	10
HD36112	A5IVe	5 27	25		440	17	2	150	140	60	15	3	250	>0.5
HD233517	K2	8 19	53	48	≤ 30			40(??)	20	54	60		<1	
TW Hya	K7Ve	10 59	-34	14	1,200	12	0.6	100(?)	140	20	0	3	20	20
HD98800	K6V	11 19	-24	13	<15			50(?)	20	55	60	<10	<0.5	0.3
HR4796	A0V	12 33	-39	6	<80			75	50	100	60	3	<7	0.2
HD135344	F4Ve	15 12	-36		400	3	2.2	100	100	38	15		6-140	30
HD141569	A0Ve	15 47	-3	-6	130	-8	7.6	200	130	60	60	<10	20-460	>0.1
HD143006	G5V	15 55	-22	-1	120	-2	3.4	57	120	20	35	<10	1	5
HD144432	A9/FOV	16 03	-27		<80			100	100	43	60	<10	<2	20
V718 Sco	A8III/IV	16 10	-22		120	-2	5	200	140	46	45	<10	30-900	15
51 Oph	B9.5Ve	17 28	-23	-12	<80			70	140	130	60	3	<6	

Symbols used: RA, right ascension; dec., declination; $V_{\odot}(\text{opt.})$, radial velocity with respect to the Sun from optical spectrum; $T_{\text{B}}(^{12}\text{CO})$, main beam brightness temperature for the ^{12}CO $J=2-1$ transition. In addition, ^{13}CO was detected at HD34700, HD36112 and TW Hya with $T_{\text{B}}=40$, 60 and 300 mK, respectively, and upper limits to the ^{13}CO intensity were obtained at HD135344, HD141569, HD143006 and V718 Sco with $T_{\text{B}} < 100$, < 30 , < 30 and < 40 mK, respectively. $V_{\odot}(\text{CO})$, radial velocity with respect to the Sun of the CO $J=2-1$ line; ΔV , full width at half maximum power of the $J=2-1$ line; D , estimated distance to a star, often quite uncertain; R , assumed or derived outer radius of molecular disk, usually quite uncertain given our limited data to date; T_{CO} (at R), excitation temperature of the CO rotational ladder at the outer edge of the molecular disk assuming that disk temperature is proportional to distance from the central star to the -0.5 power. Assumed incl., assumed disk inclination (face-on = 0°). 60° is the most probable inclination of a random sample, but face-on disks are more easily detected. Derived disk masses are not sensitive to the value assumed for the inclination. Age, estimated stellar age; M_{H_2} , total mass of gas in orbiting disk in Earth masses (318 Earth masses = 1 Jupiter mass) between R_i and R . For HD135344, HD141569 and V718 Sco the lower bound on the range of tabulated gas masses obtains if the detected ^{12}CO line is optically thin and the upper bound is set by the upper limit to the ^{13}CO line intensity. M_{d} , total mass of dust particles, with radii between a few micrometres and a few centimetres.

A portion of the data is summarized in Table 1 and Fig. 1. We assume that the ^{12}CO and ^{13}CO emission regions have the same radius R (Table 1) and that the $^{12}\text{CO}/^{13}\text{CO}$ abundance ratio is 89, and derive the CO optical depths and surface (column) densities (cm⁻²) that characterize the outer portions of the molecular disks^{24,25}. The total mass of CO, in the region where giant planets might form, is obtained by integrating these surface densities between R and an inner radius R_i which is set by the condensation temperature of water ice²⁶. If R_i is proportional to the square root of stellar luminosity then, for a pre-main-sequence late K-type star such as HD98800 or TW Hydrae, $R_i = 4$ AU, and for a main-sequence A0 star, $R_i = 30$ AU. We assume that the disk surface density is proportional to distance from the central star to the -1.5 power²⁷. Reasonable uncertainties in this exponent of ± 0.5 (ref. 27) alter the derived CO mass between R and R_i by a factor of < 1.5 . Then, assuming the standard interstellar ratio of number of H₂ and CO molecules (10^4), we calculate the H₂ masses given in Table 1.

Because of significant uncertainties in the age of, and distance to, many of the stars in Table 1, it is not now possible to make a definitive statement about orbiting gas mass as a function of stellar age. Even so the H₂ masses in Table 1 suggest that a few million years after a star is born the mass of remaining H₂ is typically much less than the mass of Jupiter (but see ref. 11). Usually, this decline of gas mass near young stars is more rapid than the decline in the mass of orbiting particulate matter (see the discussion of relative gas and dust masses below).

The primary method in astronomy for determination of the age of a young star is comparison of its position on a Hertzsprung-Russell diagram with calculations of evolutionary models (for example, ref. 28). For the sample in Table 1 this is possible, at present, only for the four stars HR4796 and HD32509, HD35187 and HD36112, whose distances are known. Often a good secondary technique is examination of Balmer emission lines. H α lines which are broad, strongly in emission and strangely shaped are usually characteristic of 'T Tauri' and Herbig Ae stars whose ages are estimated, via the primary technique, to be $< 10 \times 10^6$ yr. In contrast, for the G dwarfs in the Pleiades cluster, age 70×10^6 yr, H α is not seen in overt emission, that is "no portion of the line profile rises above the continuum"²⁹. In our Hamilton echelle spectra from the 3-m telescope at Lick Observatory³⁰, the G-type stars HD34700 and

HD143006 display H α lines in emission. Our spectra also reveal an extremely strong H α emission line at TW Hya, and strong H α emission lines with complicated shapes at HD141569 and V718 Scorpii³⁰.

Ages for some other stars in Table 1 are less well constrained. The age of 49 Ceti is estimated by analogy with HR4796²¹. HD98800 is estimated to be less than 10×10^6 yr old^{18,20} and our echelle spectra³⁰ show Ca II H and K lines strongly in emission, in support of a young age. The otherwise undistinguished, isolated, high-latitude, K-type star HD233517 is apparently a strong IRAS source³¹. With the Hamilton echelle we found that it has a deep lithium line at 6,707 Å, consistent with a very young age if it is a pre-main-sequence star (which is not yet established; note the large heliocentric velocity listed in Table 1).

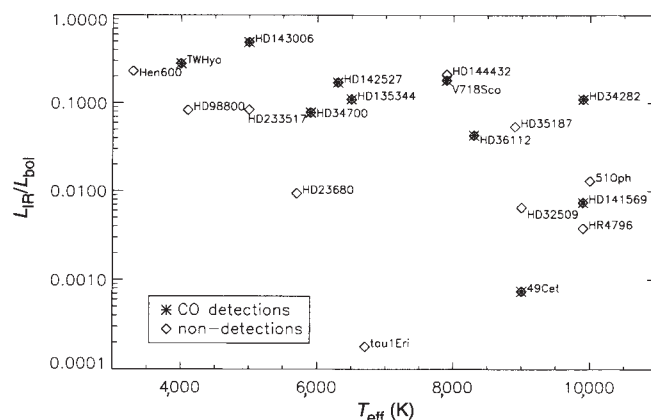


FIG. 2 Characteristics of stars studied in this work. Abscissa, the effective temperature of the target star. Ordinate, the infrared luminosity L_{IR} of a star divided by the total luminosity of the star, L_{bol} ; this is a measure of the percentage of the stellar energy that is absorbed by surrounding dust particles. This sample is very dusty in comparison to the zodiacal dust out to ~ 5 AU in our Solar System, for which $\log(L_{\text{IR}}/L_{\text{bol}}) = -7$. The data shown here suggest that the CO detection rate, and hence the CO-to-dust ratio, is not a dramatic function of the temperature of the central star as it might be if, for example, CO were readily photodissociated near A-type stars.

The detected CO lines do not obviously show the double-peaked profiles predicted by models of orbiting disks^{24,25}. There are various explanations for this, such as low signal-to-noise in our data at high spectral resolution, emission from low-velocity gas at large distances from the stars (which fills in the centre of the line profile), and large turbulent velocities in protoplanetary disks. The last explanation is not likely if the very narrow CO profile at TW Hya (Fig. 1) characterizes the scale of such turbulent velocities.

Most previous studies of the material near young stars have focused on the mass of dust particles (for example, refs 32–34). These studies suggest that the rate of dust dissipation depends on the distance between the dust and the central star and, also, the presence or absence of a companion star. But one cannot use such observations to infer the rate of dissipation of gas. Our results imply that, typically, the gas masses are lower than would be predicted on the basis of the dust observations, if one were to assume a gas-to-dust ratio equal to that in the interstellar medium (100). For most of the stars listed in Table 1, particulate masses may be calculated from 800 or 1,100 μm continuum fluxes given in the literature (see ref. 30 for additional details). Typically, the mass carried by particles with radii between a few micrometres and a few centimetres is comparable to the gas mass (see the last two columns in Table 1). So, relative to the standard ratio of gas-to-dust mass in the interstellar medium, the gas is often depleted by as much as a factor of 100. This also appears to be the situation at β Pictoris, an A-type star that is probably 10×10^6 yr old²¹. Estimated masses of dust particles and atomic and molecular gas orbiting β Pic are of the order of an Earth mass or less^{35–37}.

For the entire sample in Table 1, the dust temperatures derived by fitting the IRAS far-infrared spectra are comparable to (or a factor of ~ 2 larger than) the likely gas temperatures listed at R (at the outer edge of a molecular disk). So the dust and gas are well mixed and the dust should effectively shield the CO against photodissociation (see also Fig. 2). A different mechanism that might deplete CO relative to H_2 is freeze-out onto cold grains. But various infrared observations (refs 38, 39 and references therein) imply that most of the CO is likely to remain in the gas phase, even in regions that are colder than the coldest listed in Table 1. It therefore seems unlikely that CO freeze-out can introduce a substantial error into our analysis. $\square$

27. Lissauer, J. J. *Icarus* **69**, 249–265 (1987).
28. D'Antona, F. & Mazzitelli, I. *Astrophys. J. Suppl. Ser.* **90**, 467–500 (1994).
29. Soderblom, D. R., Stauffer, J. R., Hudon, J. D. & Jones, B. F. *Astrophys. J. Suppl. Ser.* **85**, 315–346 (1993).
30. Zuckerman, B. in *Circumstellar Dust Disks and Planet Formation* (eds Ferlet, R. & Vidal-Madjar, A.) (Editions Frontieres, Gif sur Yvette, in the press).
31. Stencl, R. E. & Backman, D. E. *Astrophys. J. Suppl. Ser.* **75**, 905–924 (1991).
32. Beckwith, S. V. W. & Sargent, A. I. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J. I.) 521–541 (Univ. Arizona, Tucson, 1993).
33. Jensen, E. L. N., Mathieu, R. D. & Fuller, G. A. *Astrophys. J.* **429**, L29–L32 (1994).
34. Strom, S. E. *Revista Mexicana de Astronomia y Astrofisica* **29**, 23–29 (1994).
35. Hobbs, L. M., Vidal-Madjar, A., Ferlet, R., Albert, C. E. & Gry, C. *Astrophys. J.* **293**, L29–L33 (1985).
36. Zuckerman, B. & Becklin, E. E. *Astrophys. J.* **414**, 793–802 (1993).
37. Savodini, M. & Galletta, G. *Astr. Astrophys.* **285**, 467–468 (1994).
38. Tielens, A. G. G. M., Tokunaga, A. T., Geballe, T. R. & Bass, F. *Astrophys. J.* **381**, 181–199 (1991).
39. Chiar, J. E., Adamson, A. J., Kerr, T. H. & Whittet, D. C. B. *Astrophys. J.* **426**, 240–248 (1994).
40. Herbst, W. & Koret, D. L. *Astr. J.* **96**, 1949–1955 (1988).
41. Franchini, M., Covino, E., Stalio, R., Terranegra, L. & Chavarria, K. C. *Astr. Astrophys.* **256**, 525–532 (1992).
42. Wolszczan, A. & Frail, D. A. *Nature* **355**, 145–147 (1992).

ACKNOWLEDGEMENTS. We thank C. Kahane for obtaining some initial CO data that gave early momentum to this project, D. Popper, S. Balm, M. Jura and G. Schultz for obtaining echelle spectra of some of the stars in Table 1, and G. Schultz and R. White for reducing these spectra. We also thank N. Houk for supplying preliminary (unpublished) spectral types and/or luminosity classes for HD 32509, 34700, 35187 and 36112. This research was funded in part by NASA.

Survival of isotopically heterogeneous graphite in a differentiated meteorite

Ahmed El Goresy*, Ernst Zinner† & Kurt Marti‡

* Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany

† McDonnell Center for the Space Sciences and the Physics Department, Washington University, St Louis, Missouri 63130, USA

‡ Department of Chemistry, 0317, University of California at San Diego, La Jolla, California 92093, USA

PRIMITIVE meteorites (carbonaceous chondrites and unequilibrated ordinary chondrites) are isotopically heterogeneous, indicating that the material from which the Solar System formed was not completely homogenized^{1–4}. On the other hand, isotopically heterogeneous material is not expected to survive in thermally 'processed' planetary or asteroidal objects. The achondrite meteorite Acapulco is a remnant of one such object; its petrographic and trace-element characteristics suggest that the parent body experienced pervasive heating and partial melting⁵. Here we report the discovery of graphite grains in the Acapulco meteorite that have a wide range of carbon and nitrogen isotopic compositions ($\delta^{13}\text{C}$ ranging from -34 to -8‰ and $\delta^{15}\text{N}$ from -154 to -67‰). The graphite is associated with metal, and in some cases, graphite grains associated with the same metal grain have very different isotopic compositions. These findings suggest that the graphite grains retain the isotopic signatures of a diverse range of precursor materials, despite the high temperatures reached in the parent asteroid.

The Acapulco meteorite belongs to a recently recognized small group of primitive achondrites: the Acapulcoites. Acapulco is considered a primitive achondrite because of its pronounced chemical affinities to H-chondrites^{5,6} but Acapulcoites originated from an asteroid with a distinct oxygen-isotope composition⁷. Many Acapulcoites, including Acapulco itself, have experienced heating, leading to partial melting and recrystallization of their constituents. Petrological and trace-element analyses of Acapulco have led to an estimate of 20% partial melting at equilibration temperatures close to or slightly above $1,200^\circ\text{C}$ (ref. 5). However, the ^{207}Pb – ^{206}Pb age of Acapulco phosphates (4.557 ± 0.002 Gyr)⁸ indicates fast cooling and an early closure of the U–Pb system soon after crystallization. ^{39}Ar – ^{40}Ar

Received 28 June; accepted 14 December 1994.

1. Ruden, S. P. & Pollack, J. B. *Astrophys. J.* **375**, 740–760 (1991).
2. Podolak, M., Hubbard, W. B. & Pollack, J. B. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J. I.) 1109–1147 (Univ. Arizona, Tucson, 1993).
3. Lissauer, J. J. *Rev. Astr. Astrophys.* **31**, 129–174 (1993).
4. Wetherill, G. W. in *Planetary Systems: Formation, Evolution, and Detection* (ed. Roettger, E.) (Kluwer Academic, Dordrecht, in the press).
5. Lissauer, J. J., Pollack, J. B., Wetherill, G. W. & Stevenson, D. J. in *Neptune* (ed. Cruikshank, D.) (Univ. Arizona, Tucson, in the press).
6. Lin, D. N. C. & Papaloizou, J. in *Protostars & Planets II* (eds Black, D. C. & Matthews, M. S.) 981–1072 (Univ. Arizona, Tucson, 1985).
7. Adams, F. C. & Lin, D. N. C. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J. I.) 721–748 (Univ. Arizona, Tucson, 1993).
8. Sargent, A. I. & Beckwith, S. V. W. *Astrophys. J.* **382**, L31–L35 (1991).
9. Weintraub, D. A., Masson, C. R. & Zuckerman, B. *Astrophys. J.* **344**, 915–924 (1989).
10. Kawabe, R., Ishiguro, M., Omodaka, T., Kitamura, Y. & Miyama, S. M. *Astrophys. J.* **404**, L63–L66 (1993).
11. Koerner, D. W., Sargent, A. I. & Beckwith, S. V. W. *Icarus* **106**, 2–10 (1993).
12. Koerner, D. W., Sargent, A. I. & Beckwith, S. V. W. *Astrophys. J.* **408**, L93–L96 (1993).
13. Dutrey, A., Guilloteau, S. & Simon, M. *Astr. Astrophys.* **286**, 149–159 (1994).
14. Skrutskie, M. F. et al. *Astr. J.* **102**, 1749–1752 (1991).
15. Yamashita, T. et al. *Astrophys. J.* **402**, L65–L67 (1993).
16. Wetherill, G. W. *Science* **253**, 535–538 (1991).
17. Oudmaier, R. D. et al. *Astr. Astrophys. Suppl.* **96**, 625–643 (1992).
18. Gregorio-Hetem, J., Lépine, J. R. D., Quast, G. R., Torres, C. A. O. & de la Reza, R. *Astr. J.* **103**, 549–563 (1992).
19. Zuckerman, B. & Becklin, E. E. *Astrophys. J.* **406**, L25–L28 (1993).
20. Fekel, F. C. & Bopp, B. W. *Astrophys. J.* **419**, L89–L92 (1993).
21. Jura, M., Zuckerman, B., Becklin, E. E. & Smith, R. C. *Astrophys. J.* **418**, L37–L40 (1993).
22. Prinn, R. G. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J. I.) 1005–1028 (Univ. Arizona, Tucson, 1993).
23. Sargent, A. I. & Welch, W. J. *Rev. Astr. Astrophys.* **31**, 297–343 (1993).
24. Omodaka, T., Kitamura, Y. & Kawazoe, E. *Astrophys. J.* **396**, L87–L90 (1992).
25. Beckwith, S. V. W. & Sargent, A. I. *Astrophys. J.* **402**, 280–291 (1993).
26. Stevenson, D. J. & Lunine, J. I. *Icarus* **75**, 146–155 (1988).

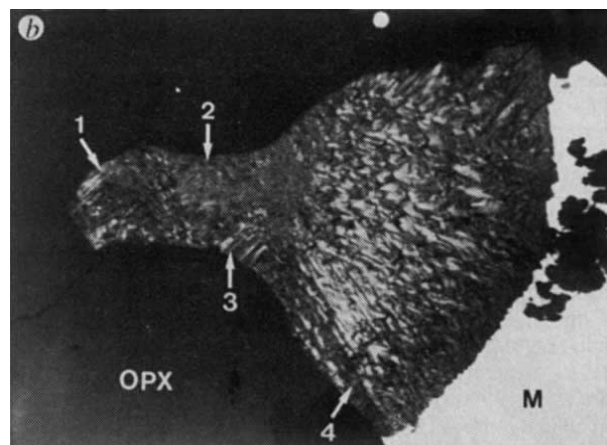
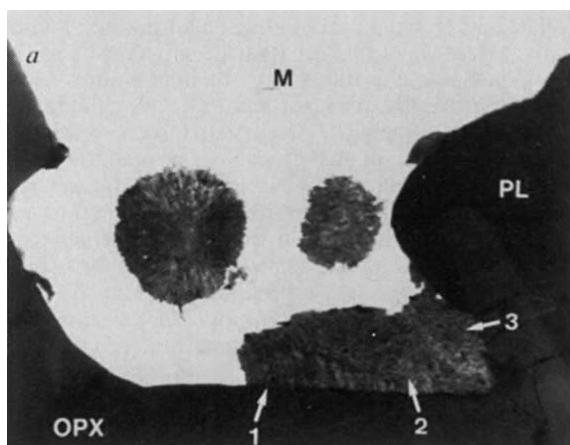


FIG. 1 *a*, Two spherulitic graphite inclusions inside a metal grain (M) from the Acapulco meteorite as observed in reflected-light microscopy. Note the radial structure of the graphite flakes expressed in the wavy variation of the grey shades resulting from bireflectance of graphite in reflected polarized light. The metal is embedded in plagioclase (pl) and orthopyroxene (opx). At the lower right is a composite grain of feathery graphite, embedded in plagioclase and orthopyroxene, that consists of three distinct segments (1, 2 and 3). Width of photograph, 90 μm . *b*, A large feathery graphite grain (160 μm in diameter)

consisting of four segments with different optical orientations of the individual feathers (1, 2, 3 and 4). These different orientations can be recognized as different shades of grey resulting from differences in bireflection. The graphite is attached to the metal (M) but is mostly embedded in orthopyroxene (opx). Sometimes it partially protrudes into the metal. Some feathers that protrude into the metal from below the section surface appear as rounded enclaves in the metal, near the curved metal-graphite interface on the right-hand side of the photograph.

analyses^{9,10} of Acapulcoites yielded undisturbed plateau ages of 4.50 and 4.514 ± 0.020 Gyr. In spite of the high temperatures experienced by Acapulco, the meteorite retained many chemical and isotopic signatures that pre-date melting of the parental material. It contains large abundances of volatile and trapped noble gases⁵, and at least three constituents with distinct nitrogen-isotope compositions: silicates ($\delta^{15}\text{N} = +15\text{‰}$), metal ($\delta^{15}\text{N} = -150\text{‰}$) and chromite ($\delta^{15}\text{N} = -75$ to -82‰)^{11,12}. It is unclear why nitrogen in metals, chromite and silicates did not isotopically equilibrate during the melting event. An important additional question is whether the metals and chromite are the only carriers of isotopically light nitrogen in Acapulco.

To address these issues, we analysed graphite and cohenite (Fe_3C) in metal grains from Acapulco for their C- and N-isotope compositions. Three polished thin sections of Acapulco were studied by reflected-light and secondary-electron microscopy. Every section contains graphite, mainly associated with metal. Graphite grains occur mostly in two distinct assemblages: some are entirely enclosed in the metal, whereas others are attached to it and intimately intergrown with, and surrounded by, silicates and in some cases by chromite. The graphite grain sizes vary from 2 to 160 μm . We can distinguish three distinct morphologies: (1) Spherulitic graphites (15–70 μm in diameter), the majority of which are entirely sealed off in FeNi metal grains with apparent preference for kamacite (α -FeNi metal). Some spherules are partially in contact with neighbouring silicates. They usually consist of radially intergrown graphite flakes originating from the centres of composite spheres (Fig. 1*a*). The radial structure can be recognized in the optical microscope in reflected polarized light, either with one polarizer or between crossed polarizers. Some spheres that were not sectioned through their centres display a granular texture that expresses the cross-sections of individual graphite flakes. Some of the metal grains with spherulitic graphites are located close to the fusion crust. These graphite spheres are enclosed in a eutectic intergrowth of FeNi-FeS, indicating melting at a temperature $T \geq 990^\circ\text{C}$. (2) Multicomponent graphite inclusions (50–160 μm in diameter) intergrown with both silicates (mostly plagioclase and orthopyroxene) and metal. Every inclusion consists of several segments of fan-like feathery structures with different orientations of the individual feathers (Fig. 1*b*). Feathers of the fan-like segments radiate from several points, either at the boundary with

a neighbouring graphite segment or with silicate but not with metal. Every graphite inclusion is, however, attached to a metal grain, even if it is mainly enclosed in and intimately intergrown with silicate (Fig. 1*b*). Both of these graphite morphologies frequently occur close to one another (as close as 15 μm) in the same metal grains (Fig. 1*a*). (3) Rare fine-grained ($\leq 1 \mu\text{m}$) graphite filigree, consisting of fine threads attached to each other. They occur without any apparent preferred crystallographic orientations in the metal.

The metal/graphite textural relationship of the graphites excludes their formation by exsolution from FeNi metal on cooling. No metal grains with graphite were found to contain cohenite. Only one metal grain, barren of graphite, was found to enclose several patches of cohenite and schreibersite (FeNi_3P) at the kamacite-taenite (γ -FeNi metal) boundary. This is in contrast to graphite-cohenite assemblages in IIB iron meteorites¹³ where both phases occur together.

Ion-microprobe isotopic analyses of graphite and cohenite were made with a Cs^+ beam of 2–3 μm in diameter. Nitrogen was measured as CN^- according to the procedures described in ref. 14. Carbon-isotope ratios in cohenite were calibrated by measuring cohenite and graphite standards ($\delta^{13}\text{C}_{\text{graphite}} = -6.53 \pm 0.37\text{‰}$; $\delta^{13}\text{C}_{\text{cohenite}} = -18.27 \pm 0.20\text{‰}$) from the Cañon Diablo meteorite, the C-isotope compositions of which have been determined by Pillinger and co-workers (personal communication).

All graphite grains analysed in the ion probe contain substantial amounts of N ($0.006 < \text{CN}^-/\text{C}^- < 0.035$) with the N content of feathery graphite being, on average, higher than that of spherulitic graphite (Fig. 2*a*). In contrast to graphite, cohenite has very low nitrogen contents (CN^-/C^- ratios are as low as 3×10^{-5}), which did not permit N-isotope analyses in this phase. Spherulitic and feathery graphites have distinct N- and C-isotope signatures. With the exception of three spherules, $\delta^{13}\text{C}$ values in spherulitic graphite vary between -35 and -15‰ and are correlated with $\delta^{15}\text{N}$ values (Fig. 2*b*). The slope of the correlation line for spherulitic graphite, which goes through $\delta^{13}\text{C} = \delta^{15}\text{N} \approx 0$, is 3.25. The C-isotope composition of cohenite is $\delta^{13}\text{C} = -28 \pm 3\text{‰}$. This isotopic composition is slightly lighter than the range found for cohenite in iron meteorites¹⁵. The C-isotope compositions of the different graphite morphologies in Acapulco are different from those measured in graphite in iron

meteorites ($\delta^{13}\text{C} = -5$ to -6%) and the range is much wider than that for taenite ($\delta^{13}\text{C} = -19.7$ to -22.1%) in those meteorites¹⁵. Deines and Wickman¹⁵ considered the difference in the C-isotope ratios between graphite and cohenite to be the result of equilibration between these two phases. Because we did not find any metal grains with coexisting graphite and cohenite we cannot reach a conclusion whether such a mechanism is responsible for the C-isotope composition of cohenite in Acapulco. In Acapulco the range of the graphite morphologies brackets the value for cohenite, indicating that graphite could not have resulted from the breakdown of cohenite during cooling.

The results presented here are evidence that isotopic heterogeneities on a small spatial scale are not only restricted to unequilibrated chondrites, but apparently occur also in meteorite samples of asteroids that have experienced extensive recrystallization and partial melting. In spite of the high temperature reached in the Acapulco parent body, the C and N isotopes did not equilibrate and retained some signatures of isotopically diverse precursors. We note that the N-isotope compositions of individual graphite grains bracket the range reported for Acapulco metal¹¹. It remains to be seen whether additional nitrogen is contained in the metal or whether all nitrogen that has previously been measured in association with metal^{11,12} is entirely sited in graphite. Although the feathery graphite is intimately intergrown with the silicates, its N is isotopically similar to or even lighter than that of the spherulites enclosed in the metal. Petrological studies⁵ of Acapulco suggest a short early heating event to peak temperatures close to or slightly above $1,200^\circ\text{C}$. This event resulted in 20% partial melting of the metal and silicate constituents. The time-temperature path appears to be complex, starting with a high cooling rate of $\sim 100\text{ K per }10^6\text{ yr}$ (refs 5, 8). The faster cooling rate of Acapulco compared with H-chondrites may indicate a small parent body⁵. The limited degree of partial melting was evidently not sufficient for extensive differentiation and core formation of the parent body. Despite the high degree of chemical equilibration between the coexisting silicates, metal and chromite are not uniformly distributed in different samples⁵ indicating a lack of homogenization during melting. The limited peak temperature of this event and the high cooling rate apparently prevented C- and N-isotope equilibration between the different coexisting graphite morphologies. The use of this lack of equilibration for obtaining tighter constraints on the time-temperature path of the Acapulco parent body (and the resulting implications for planetary differentiation and core formation) requires experimental data on the distribution and diffusive behaviour of C and N at high temperatures in the relevant phase assemblages. These data presently do not exist. It is hoped that our preliminary findings of isotopic heterogeneity will stimulate such studies, and ultimately lead to a better understanding of planetary differentiation not only in Acapulcoites but possibly in other differentiated meteorite classes.

At present we cannot identify any primitive (possibly presolar) carbonaceous material that could explain the C- and N-isotope compositions observed in Acapulco graphite. Interstellar graphite grains in the Murchison meteorite have, on average, isotopically heavy carbon (the mean $\delta^{13}\text{C}$ value is $+1,120\%$)¹⁶ and nitrogen (that is, positive $\delta^{15}\text{N}$ values)¹⁷. Interstellar diamonds do have light nitrogen ($\delta^{15}\text{N} = -340\%$), which, together with their $\delta^{13}\text{C}$ value^{18,19} of $\sim -34\%$, would make them a potential precursor component for the feathery graphite, if mixed in proper ratios with a component of normal N and $\delta^{13}\text{C} \geq -10\%$. But the same precursor components could not account for the compositions of the spherulitic graphite plotting along array A of Fig. 2b. Even the postulated existence of a component with $\delta^{13}\text{C} \geq -10\%$ poses a problem because most Solar System carbon has $\delta^{13}\text{C} \approx -30\%$.

An alternative would be to explain the linear array of data points (A in Fig. 2b) as being due to isotopic mass fractionation

of both C and N, but a process that would produce the required (negative) fractionation effect (that is, enrichments in the light isotopes) in C and N is difficult to obtain in nature. Excesses of ^{15}N in interplanetary dust particles^{20,21}, CR chondrites²² and spherulitic carbon aggregates from Murchison²³ have been interpreted as being due to the survival of presolar organic compounds that obtained their ^{15}N enrichments by ion-molecule reactions in molecular clouds. But C-isotope ratios in these objects are quite normal, and it is not clear whether correlated C and N fractionations can be produced in this way in the first place²⁴⁻²⁶. Furthermore, we still need to explain the depletions in ^{15}N that are seen in Acapulco graphite. Clearly, more information is needed before any identification of the precursor components can be made. □

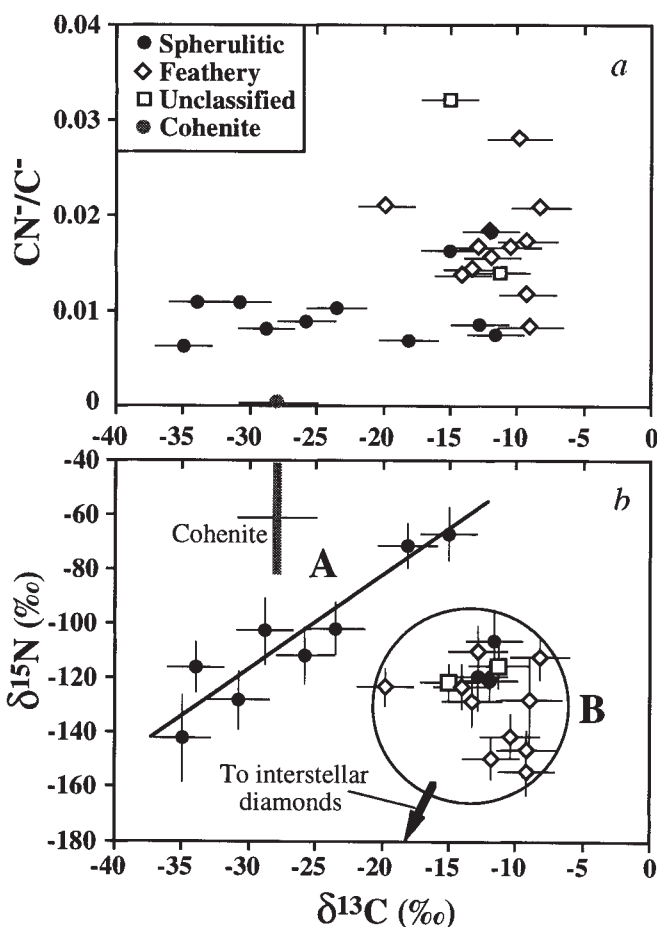


FIG. 2 CN^-/C^- ratios (from ion-probe measurements) and N- and C-isotope compositions of graphite grains (with different morphologies) and cohenite: a, CN^-/C^- versus $\delta^{13}\text{C}$ plot. The CN^-/C^- ratio is a measure of the N content of the graphite. Notice the narrow C-isotope composition of feathery graphite and the lower N content of spherulitic graphite. b, $\delta^{15}\text{N}$ versus $\delta^{13}\text{C}$ plot of spherulitic, feathery and two unidentified graphites in the Acapulco meteorite. The C-isotope composition of cohenite, the N isotopes of which could not be measured, is indicated by a vertical bar. Data points for most spherulitic graphite grains plot along a linear array (labelled A). Data points of feathery graphite plot within a limited region B (circled). Region B also contains three small grains that were classified as spherulitic graphites, based on their round shape and their enclosure in the metal, but due to poor polishing the internal radial structure of these three spherulitic graphites could not be discerned. Possibly they could also be either graphite filigree or protrusions of feathery graphite in metals. The arrow points to the isotopic composition ($\delta^{13}\text{C} = -34\%$; $\delta^{15}\text{N} = -340\%$) of interstellar diamonds present in primitive chondritic meteorites.

Received 12 September; accepted 20 December 1994.

1. Wasserburg, G. J. *Earth planet. Sci. Lett.* **86**, 129–173 (1987).
2. Clayton, R. N., Hinton, R. W. & Davis, A. M. *Phil. Trans. r. Soc. A325*, 483–501 (1988).
3. Lee, T. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 1063–1089 (Univ. Arizona Press, Tucson, 1988).
4. Anders, E. & Zinner, E. *Meteoritics* **28**, 490–514 (1993).
5. Zipfel, J., Palme, H., Kennedy, A. K. & Hutcheon, I. D. *Geochim. cosmochim. Acta* (in the press).
6. Palme, H. et al. *Geochim. cosmochim. Acta* **45**, 727–752 (1981).
7. Clayton, R. N., Mayeda, T. K. & Nagahara, H. (abstr.) *Lunar planet. Sci.* **XXIII**, 231–232 (1992).
8. Göpel, C., Manhès, G. & Allègre, C. J. (abstr.) *Meteoritics* **27**, 226 (1992).
9. Bogard, D. D., Garrison, D. H., McCoy, T. J. & Keil, K. (abstr.) *Lunar planet. Sci.* **XXIV**, 141–142 (1993).
10. Pellas, P., Trierloff, M. & Feini, C. *Geochim. cosmochim. Acta* (submitted).
11. Sturgeon, G. & Marti, K. *Proc. lunar planet. Sci. Conf.* **21**, 523–525 (1991).
12. Kim, Y. & Marti, K. (abstr.) *Lunar planet. Sci.* **XXV**, 703–704 (1994).
13. El Goresy, A. *Geochim. cosmochim. Acta* **29**, 1131–1151 (1965).
14. Zinner, E. K., Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **53**, 3273–3290 (1989).
15. Deines, P. & Wickman, E. *Geochim. cosmochim. Acta* **39**, 547–558 (1975).
16. Anders, E. & Zinner, E. *Icarus* **112**, 303–309 (1994).
17. Amari, S., Hoppe, P., Zinner, E. & Lewis, R. S. *Nature* **365**, 806–809 (1993).
18. Virag, A., Zinner, E., Lewis, R. S. & Tang, M. (abstr.) *Lunar planet. Sci.* **XX**, 1158–1159 (1989).
19. Russell, S. S., Arden, J. W. & Pillinger, C. T. *Science* **254**, 1188–1191 (1991).
20. Stadermann, F. J., Walker, R. M. & Zinner, E. (abstr.) *Meteoritics* **24**, 327 (1989).
21. Stadermann, F. J. thesis, Univ. Heidelberg (1991).
22. Ash, R. D., Morse, A. D. & Pillinger, C. T. (abstr.) *Meteoritics* **28**, 318–319 (1993).
23. Zinner, E., Amari, S., Wopenka, B. & Lewis, R. S. *Meteoritics* (in the press).
24. Adams, N. G. & Smith, D. *Astrophys. J.* **247**, L123–L125 (1981).
25. Langer, W. D., Graedel, T. E., Frerking, M. A. & Armentrout, P. B. *Astrophys. J.* **277**, 581–604 (1984).
26. Langer, W. D. in *Astrochemistry of Cosmic Phenomena* (ed. Singh, P. D.) 193–197 (Kluwer Academic, Dordrecht, 1992).

ACKNOWLEDGEMENTS. A.E. thanks R. M. Walker and the McDonnell Center for the Space Sciences for their hospitality and support during repeated visits to Washington University. We thank C. Pillinger and M. Grady for analysis of the standards, and I. Hutcheon and R. Hutchison for reviews. A.E. was supported by the Deutsche Forschungsgemeinschaft and E.Z. was supported by NASA.

Atomic-layer titration of surface reactions

R. M. Tromp & T. Michely

IBM Research Division, T. J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA

THE composition of surface layers on solid substrates is critical to many technological problems, such as semiconductor processing and heterogeneous catalysis. Deposited adatoms commonly react with the surface of the substrate to form a new, often two-dimensional structure. The stoichiometry of the surface phase can be difficult to determine by techniques of surface analysis, however; one can monitor the quantity of material deposited (using a quartz microbalance, for example), but the amount of substrate consumed is harder to assess. Here we present a technique for monitoring surface reactions in real time, which allows us to determine the amount of substrate material incorporated into the surface phase. We use the fact that surface reactions commonly involve surface roughening owing to extraction of substrate atoms. By supplying such atoms in advance in the form of sub-monolayer islands, we can titrate the subsequent adatoms against these pre-deposited islands of 'substrate' until a perfectly smooth surface (observed microscopically) is obtained. This should allow accurate characterization of surface reactions with stoichiometric ratios as great as 10:1.

As a metal is adsorbed on a semiconductor surface, it attacks the substrate and reacts with it. This generally leads to a roughening of the surface, with the reaction proceeding predominantly at or near atomic steps. The rough morphology that results is not conducive to quantitative analysis. But the situation can easily be reversed by simple modification of the surface before metal adsorption. Suppose that half a monolayer of substrate atoms (say, Si) are consumed during the formation of a given metal-induced structure. Rather than waiting for the metal atoms to extract this material from the surface, we can adsorb

half a monolayer of Si on the surface (from an evaporator) before metal adsorption. If done appropriately, this half monolayer will be present on the surface in the form of numerous small two-dimensional islands, giving rise to a rough surface, dense in atomic steps. When metal atoms are adsorbed on this surface they will attack these steps and consume the adsorbed Si islands. If the correct amount of Si was pre-adsorbed, the surface will be atomically smooth when the reaction is completed. Alternatively, one may follow the evolution of a single isolated island during the reaction. For instance, expansion of the island during reaction indicates that the substrate material is more dilute in the new phase than in the starting surface. We may think of this as a titration experiment in which the mass balance of the surface reaction is monitored by the evolving atomic-layer morphology of the surface. We therefore refer to this method as atomic-layer titration.

In Fig. 1a we show an image of a clean Si(001) surface, obtained with low-energy electron microscopy^{1,2} (LEEM). The surface undergoes a (2×1) dimer reconstruction. Upon crossing an atomic step, the dimer orientation rotates by 90° . In the dark-

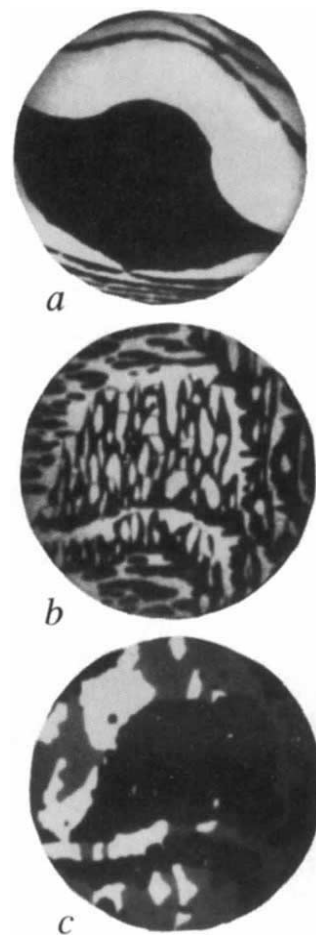


FIG. 1 a, Dark-field low-energy electron microscopy (LEEM) image of a clean Si(001) (2×1) surface. Atomic terraces with orthogonal (2×1) and (1×2) dimer orientations are shown in black and white, respectively, and are separated by atomic steps 1.36 Å high. b, Si(001) surface on which half a monolayer of Si has been deposited before Ag adsorption. Islands, many atomic layers high, can be seen all over the surface. c, After adsorption of 0.5 monolayers of Ag the surface is atomically smooth, and the Si islands seen in b have disappeared owing to reaction with Ag. The grey-white contrast on the highlighted terraces is due to a lack of mirror symmetry in the new (2×3) Ag unit cell, giving rise to two domain orientations on each terrace (showing up in grey and white contrast). Field of view is 4 µm. (The area shown in a is similar to, but not the same as, that in b and c.)

field imaging conditions employed here, terraces with one dimer orientation appear bright, while terraces with the orthogonal dimer orientation appear black. The alternation between black and white terraces is a direct manifestation of the presence and locations of surface steps³. When this surface is exposed to silver, at a temperature of 550 °C, it transforms to a well ordered (2×3) Ag structure, the unit cell of which is six times larger than that of an ideally terminated, unreconstructed surface⁴. If we allow this reaction to proceed on the surface shown in Fig. 1a, the final surface is rough because of reaction of Ag with the substrate. In Fig. 1b a surface is shown on which 0.5 monolayers of Si was adsorbed at room temperature, and subsequently heated to 550 °C. Many small two-dimensional islands can be seen. When exposed to Ag, the reaction is initiated at the edges of the islands which are consumed to form the (2×3) structure. When formation of the (2×3) phase is completed (Fig. 1c) the surface is atomically smooth. If less than 0.5 monolayers of Si is pre-adsorbed, the surface contains many atomic-layer holes; if more Si is pre-adsorbed, the islands are not fully consumed. When the Ag is thermally desorbed, the two-dimensional Si islands reappear. From a quantitative analysis of these images, we have determined that 0.52 ± 0.04 monolayers of Si are consumed in this reaction. The amount of Ag was measured as 0.54 ± 0.05 monolayers, using medium-energy ion scattering. So each (2×3) unit cell contains three Si and three Ag atoms. (The grey-white contrast of the highlighted areas in Fig. 1c indicates a lack of mirror-symmetry of the (2×3) unit cell, giving rise to two equivalent domain orientations on each terrace⁵.)

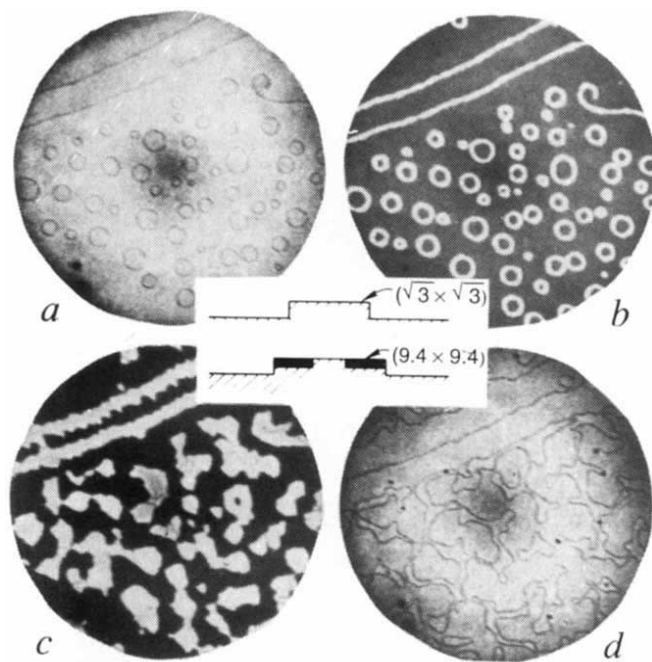


FIG. 2 a, Bright-field LEEM image of the Si(111) ($\sqrt{3} \times \sqrt{3}$) Al surface with pre-deposited two-dimensional Si islands. As the surface has the same phase everywhere, the contrast is uniform. Surface steps show up as dark lines. b, After adsorption of additional Al, the (9.4×9.4) phase starts to form at the step edge. The central ($\sqrt{3} \times \sqrt{3}$) regions of the islands shrink, while the total island areas expand, indicating outward mass transport of Si from the ($\sqrt{3} \times \sqrt{3}$) phase into the (9.4×9.4) phase. c, d, Further formation of the (9.4×9.4) phase, until completion. Because insufficient Si was pre-deposited to exactly balance the demand, new steps (two-dimensional holes) are formed on the surface after the pre-deposited islands are consumed. Inset, schematic illustration of the partial conversion of a ($\sqrt{3} \times \sqrt{3}$) island to the (9.4×9.4) structure. The mass balance can be determined from measurements of the island areas in a and b, or from a direct comparison of the surface morphologies in a and d. Field of view is 5 μm .

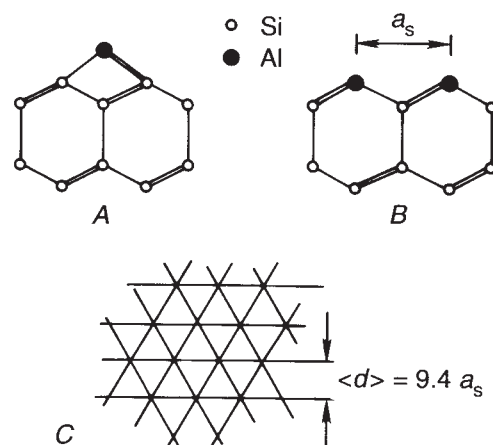


FIG. 3 A, In the Si(111) ($\sqrt{3} \times \sqrt{3}$) Al surface, Al adatoms occupy adatom sites on top of the outer Si(111) double layer. B, In the Si(111) (9.4×9.4) Al surface, Al atoms occupy substitutional sites in the top of the outer Si(111) double layer, displacing 1 monolayer of Si. C, Owing to large compressive stress in the substitutional Al structure, the surface contains a high density of misfit dislocations, giving rise to the incommensurate (9.4×9.4) diffraction pattern. $\langle D \rangle$ is the average dislocation spacing, a_s is the surface lattice spacing.

Similarly, we have studied the adsorption of Al on Si(111). At one-third of a monolayer of Al a ($\sqrt{3} \times \sqrt{3}$) structure forms, which transforms into an incommensurate structure at higher coverage^{6,7}. Figure 2a shows an image of the ($\sqrt{3} \times \sqrt{3}$) Al surface, with Al atoms adsorbed in the form of adatoms (Fig. 3A). Again, pre-adsorbed Si islands, as well as atomic steps, can be seen. When the Al coverage is increased, the island edges are attacked (Fig. 2b). During formation of the incommensurate (9.4×9.4) phase, the central dark ($\sqrt{3} \times \sqrt{3}$) regions of the two-dimensional island shrink, while the peripheral bright (9.4×9.4) regions expand beyond the original island size, indicating Si mass transport from the ($\sqrt{3} \times \sqrt{3}$) phase to the (9.4×9.4) phase. From the ratio of the original island size to that of a partially or fully reacted island we can derive the density of Si atoms in the (9.4×9.4) phase. Figure 2c and d show the evolution to a complete (9.4×9.4) surface. In this case, because insufficient Si was pre-deposited, the final surface is not perfectly smooth. After the pre-deposited Si was consumed, reaction between Si and Al started between the islands, releasing Si and creating additional atomic steps, apparent from the complex step arrangement in Fig. 2d. Because surface steps are easily resolved, the mass transport between the morphologies of Fig. 2a and d can be determined quantitatively. With both methods (that is, evolution of a single island, or evolution of the full surface morphology) we find that 0.52 ± 0.03 of a double layer, or close to a monolayer of Si is consumed in this reaction. Independently, we determine that 0.68 monolayer of Al is present in the (9.4×9.4) structure. Addition of 1 monolayer of Si to the ($\sqrt{3} \times \sqrt{3}$) surface results in a structure in which the outer half of the Si(111) double layer is missing. The missing top half is supplied by Al, occupying a substitutional site (Fig. 3B) as originally proposed by Lander and Morrison⁶. Naively, one would expect low-energy electron diffraction (LEED) to show a (1×1) pattern, not a (9.4×9.4) pattern. However, calculations based on 'first principles' have shown that such a surface contains a large compressive stress⁸. This stress is relieved by removing rows of Al atoms from the substitutional layer, forming a dislocation-like network (Fig. 3C) with an average dislocation spacing of 9.4 surface lattice spacings. From the (9.4×9.4) periodicity of this network we predict an Al coverage of 0.65–0.70 monolayer, in excellent agreement with the measured 0.68 monolayer.

As shown above, quantitative knowledge of the amount of Si used in these reactions is critical in obtaining the atomic structure of the final phase. When the Si(111) surface is exposed to Ag, the reaction leads to the formation of a $(\sqrt{3} \times \sqrt{3})$ superstructure. Although the symmetry of the $(\sqrt{3} \times \sqrt{3})$ Ag structure is the same as that of the $(\sqrt{3} \times \sqrt{3})$ Al surface discussed above, its atomic structure is very different. This Si(111) $(\sqrt{3} \times \sqrt{3})$ Ag surface may serve as an example of how hard it is to obtain structural information without knowledge of the mass balance of the reaction. Its structure was investigated and debated for a number of years, largely because the vital information that the outer half of the first Si double layer is missing (like in the (9.4×9.4) Al surface) was not available^{9,10}. Similar problems are encountered with the Si(001) (2×3) Ag surface, for which previously proposed structural models⁴ do not incorporate the three Si atoms per unit cell that our atomic-layer titration experiments show to be present. Although determining the mass balance of surface chemical reactions alone is of course not sufficient in general to determine the structure (in all but the simplest chemical reactions), to determine the structure without this knowledge is extremely difficult.

We believe that the simple method of atomic-layer titration presented here can be applied to a wide variety of surface chemical reactions, and is not limited to the formation of Si-metal interfaces. Although we have used LEEM to follow the evolution of surface morphology, other microscopic techniques such as scanning tunnelling microscopy or ultra-high-vacuum transmission electron microscopy may be equally suitable. Also, non-imaging methods, particularly diffraction techniques (such as X-

ray and electron diffraction) may be used. The surface roughness can be monitored using the diffracted beam profiles in appropriate out-of-phase scattering conditions.

We note that although the reactions presented here occur only in the outer atomic layer, this is not always the case in other systems, and the involvement of deeper substrate layers may complicate the analysis. In addition, if the substrate contains many defects, the chemical reactions may be dominated by these instead of by surface steps. In spite of these limitations, atomic-layer titration should be a simple and reliable technique to determine the mass balance of a wide range of surface chemical reactions. □

Received 28 October; accepted 20 December 1994.

1. Telleps, W. & Bauer, E. *Ultramicroscopy* **17**, 57–65 (1985).
2. Tromp, R. M. & Reuter, M. C., *Proc. Mater. Res. Soc. Vol. 237* (eds Liang, K. S., Anderson, M. P., Bruinsma, R. F. & Scoles, G.) 349–358 (Mater. Res. Soc., Pittsburgh, 1992).
3. Tromp, R. M. & Reuter, M. C., *Phys. Rev. Lett.* **68**, 820–822 (1992).
4. Liu, X. I., Wan, K. J. & Nogami, J., *Phys. Rev. B* **49**, 7385–7393 (1994).
5. Michely, T., Reuter, M. C., Copel, M. & Tromp, R. M., *Phys. Rev. Lett.* **73**, 2095–2098 (1994).
6. Lander, J. J. & Morrison, J., *Surf. Sci.* **2**, 553–565 (1964).
7. Nishikate, K., Murakami, K., Yoshimura, M. & Kawazu, A. *Surf. Sci.* **269** **270**, 995–999 (1992).
8. Meade, R. D. & Vanderbilt, D. *Phys. Rev. Lett.* **63**, 1404–1407 (1989).
9. Katayama, M., Williams, R. S., Kato, M., Nomura, E. & Aono, M., *Phys. Rev. Lett.* **66**, 2762–2765 (1991).
10. Ding, Y. G., Chan, C. T. & Ho, K. M., *Phys. Rev. Lett.* **67**, 1454–1457 (1991).

ACKNOWLEDGEMENTS. We thank M. Reuter and M. Copel for their assistance with the experiments. Th. M. was supported by a Feodor Lynen Research Fellowship from the Alexander von Humboldt Foundation.

Enhancement of interfacial catalysis in a biphasic system using catalyst-binding ligands

R. V. Chaudhari*, B. M. Bhanage*,
R. M. Deshpande* & H. Delmas†

* Chemical Engineering Division, National Chemical Laboratory,
Pune 411008, India

† Ecole Nationale Supérieure d'Ingénieurs en Génie Chimique,
Chemin de la Loge, 31078 Toulouse, France

To avoid the problem of separation of products from catalyst in homogeneous catalysis¹, two-phase systems have been developed in which the catalytic complex (usually a water-soluble organometallic complex) remains in one (generally aqueous) phase while the products remain in a second, immiscible phase². Catalysis relies on the transfer of organic substrates into the aqueous catalyst phase; but the limited solubility of these substrates in water leads to reaction rates much lower than those for conventional homogeneous catalysis. Here we show that catalysis at the interface of a two-phase system can be enhanced by using a 'promoter ligand' which, although soluble in the organic phase, will bind to the organometallic catalyst and thus increase its concentration close to the interface in the aqueous phase. We demonstrate this approach for the hydroformylation of 1-octene using a rhodium-based catalyst. A rate enhancement by a factor of 10–50 is observed when we introduce the promoter ligand PPh₃ in the organic phase.

Two-phase catalysis is used commercially in the Rhurchemie–Rhône Poulenc process for hydroformylation of propylene to butyraldehyde using a water-soluble rhodium triphenylphosphine trisulphonate (TPPTS) catalyst². Hydroformylation of olefins is one of the most important and largest-scale example

of homogeneously catalysed reactions, 5.9 million tonnes of oxo alcohols being produced annually by this process³. But because of the limited solubility of olefins in water, the activity of these water-soluble catalysts is considerably lower than in truly homogeneous systems. Attempts to enhance the rate of such reactions using co-solvents⁴ and supported aqueous-phase catalysis (SAPC)⁵ have been made, but the co-solvent system leads to lower selectivity by interaction of co-solvents and products, and the SAPC system can be limited by intraparticle diffusional resistance.

Our new approach involves the use of a ligand that is insoluble in the aqueous (catalyst) phase but has a strong affinity to the metal complex catalyst. Because the catalytic complex is insoluble in the organic medium, the interaction of the water-insoluble ligand (for example, PPh₃) with the water-soluble catalyst [Rh(COD)Cl]₂/TPPTS will essentially occur at the liquid–liquid interface (COD is 1,5 *cyclo*-octadiene). This is likely to enhance the concentration of the catalytic species at the interface, where it can access the reactants present in the organic phase in significantly higher concentrations. This can result in a dramatic increase in the rate of such a biphasic catalytic reaction.

We illustrate this approach for the hydroformylation of 1-octene using a water-soluble [Rh(COD)Cl]₂/TPPTS catalyst in a biphasic system. This catalyst was prepared using the procedure described in the literature⁶. The gaseous reactants H₂ and CO exist as dispersed bubbles in the system: the reaction occurs by transport of H₂ and CO to the liquid phase, followed by transport of dissolved H₂, CO and 1-octene from the organic to the aqueous phase. We quantify the catalytic activity in terms of the turnover frequency (TOF), defined as the rate of product formation per unit weight of Rh in catalyst per unit time (in kmol per kg Rh s⁻¹). The TOF values were calculated using the data for <20% conversion of 1-octene. The results are shown in Fig. 1, which clearly indicates a dramatic increase in the TOF (by a factor of 10–50) in the presence of the promoter ligand PPh₃ in the organic phase (similar results were observed for

TOF values calculated for complete conversion of 1-octene). We interpret this as an indication that the interaction between PPh_3 and the Rh-TPPTS complex at the interface increases the local concentration of the Rh complex here, where it can access the much greater concentration of 1-octene in the organic phase. (The equilibrium solubility of 1-octene in water is $2.408 \times 10^{-5} \text{ kmol m}^{-3}$ at 298 K, whereas the concentration in the organic phase is 0.85 kmol m^{-3} .) The results in Fig. 1 represent initial catalytic activity, but nearly 100% conversion of 1-octene was achieved in all the experiments under conditions of interfacial catalysis. The TOF values for conditions of interfacial catalysis are in a range of 5.11×10^{-4} to $2.4 \times 10^{-3} \text{ kmol per kg Rh s}^{-1}$, which is equivalent to a turnover of 180–800 (h^{-1}). These are comparable to the range desired (>200 turnovers h^{-1}) for industrial catalytic processes⁷.

To establish that the observed increase in the rate is indeed due to promotion of interfacial catalytic reaction, several test experiments were carried out. (1) The aqueous catalytic phase was recycled along with fresh organic phase without addition of PPh_3 . The rate was very low, comparable to that of biphasic catalytic reaction in the absence of PPh_3 (Fig. 1a). The same reaction was continued with addition of PPh_3 in the organic phase, whereupon the rate increased markedly (Fig. 1d). (2) Hydroformylation of 1-octene in a homogeneous methanol phase was studied using first, $[\text{Rh}(\text{COD})\text{Cl}]_2 + \text{TPPTS}$ (1:6) and second, $[\text{Rh}(\text{COD})\text{Cl}]_2 + \text{TPPTS} + \text{PPh}_3$ (1:6:0.2) catalysts at 373 K and $4.14 \times 10^3 \text{ kPa}$. The TOF values observed were 0.0053 and $0.0061 \text{ kmol per kg Rh s}^{-1}$, respectively. These results indicate clearly that the Rh-TPPTS- PPh_3 complex does not have a significantly higher catalytic activity than Rh-TPPTS complex. The overall higher TOF values in these experiments were mainly due to very high solubility of 1-octene in methanol compared to that of water. (3) A complex of the type $\text{HRh}(\text{CO})(\text{TPPTS})_{3-x}(\text{PPh}_3)_x$ was prepared by refluxing an aqueous solution of $\text{HRh}(\text{CO})(\text{TPPTS})_3$ (0.055 mmol in $2 \times 10^{-6} \text{ m}^3 \text{ H}_2\text{O}$) and a solution of PPh_3 (0.055 mmol) in dichloromethane ($2 \times 10^{-6} \text{ m}^3$) under an atmosphere of argon (as suggested by Borowski *et al.*⁸).

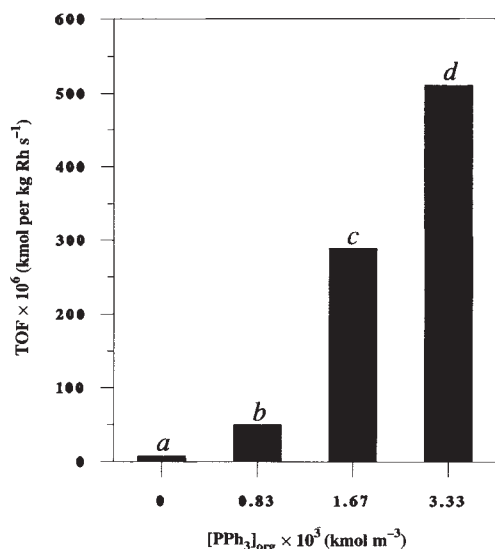


FIG. 1 Hydroformylation of 1-octene in a biphasic media. TOF, turnover frequency (see text); subscript org indicates organic-phase concentration. Reaction conditions: $[\text{Rh}(\text{COD})\text{Cl}]_2$, 0.01 kmol m^{-3} ; [1-octene], 0.85 kmol m^{-3} ; temperature, 373 K; stirring speed, 15 r.p.s.; pressure, $4.14 \times 10^3 \text{ kPa}$ ($\text{CO}/\text{H}_2=1:1$); Rh:TPPTS=1:6; aqueous phase hold-up (the ratio of the aqueous-phase volume to the total volume of the two liquid phases), 0.4; solvent system, toluene–water; total volume of liquids, $2.5 \times 10^{-5} \text{ m}^3$.

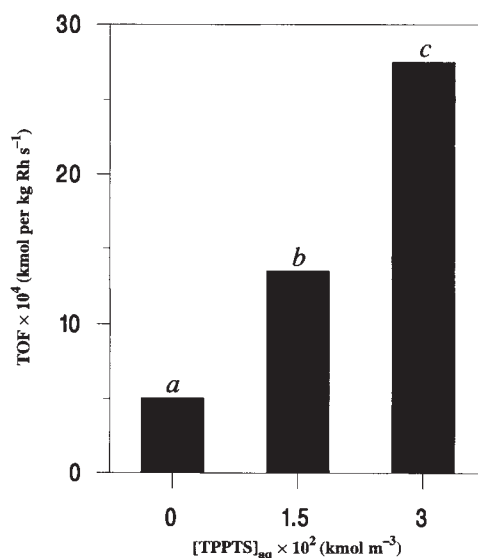


FIG. 2 Hydroformylation of allyl alcohol in a biphasic media. $[\text{TPPTS}]_{\text{aq}}$ indicates aqueous-phase concentration of triphenylphosphine trisulphonate. Reaction conditions: $[\text{HRh}(\text{CO})(\text{PPh}_3)_3]$, $5 \times 10^{-3} \text{ kmol m}^{-3}$; [allyl alcohol], 2.94 kmol m^{-3} ; temperature, 313 K; stirring speed, 15 r.p.s.; pressure $4.14 \times 10^3 \text{ kPa}$ ($\text{CO}/\text{H}_2=1:1$); aqueous phase hold-up, 0.5; solvent system, toluene–water; total volume of liquids, $2.5 \times 10^{-5} \text{ m}^3$.

*al.*⁸). The exchanged complex was precipitated from the aqueous solution by adding absolute ethanol. The ^{31}P NMR of these complexes showed the following. For $\text{HRh}(\text{CO})(\text{TPPTS})_3$: ^{31}P { ^1H } NMR (D_2O , 200 MHz, 25 °C); δ 43.6 (doublet, $J(\text{PRh})=156 \text{ Hz}$) p.p.m. For $\text{HRh}(\text{CO})(\text{TPPTS})_{3-x}(\text{PPh}_3)_x$: ^{31}P { ^1H } NMR (D_2O , 200 MHz, 25 °C); δ 43.53 (doublet, $J(\text{PRh})=156 \text{ Hz}$) p.p.m.; δ 43.12 (doublet, $J(\text{PRh})=156 \text{ Hz}$) p.p.m. The ^{31}P chemical shift for $\text{HRh}(\text{CO})(\text{TPPTS})_3$ agrees with that in previous studies⁹. The observation of an additional doublet at 43.12 p.p.m. in the exchanged complex ($\text{HRh}(\text{CO})(\text{TPPTS})_{3-x}(\text{PPh}_3)_x$) indicates that partial exchange of TPPTS has occurred by a ligand with higher electron density, such as PPh_3 .

The activity of this exchanged complex was tested for hydroformylation of 1-octene in a biphasic medium under the conditions of the experiment represented in Fig. 1d. This catalyst showed a TOF value of $2.62 \times 10^{-5} \text{ kmol per kg Rh s}^{-1}$, which is higher than that for the biphasic Rh-TPPTS complex, $7.85 \times 10^{-6} \text{ kmol per kg Rh s}^{-1}$, but much lower than that for the experiment in Fig. 1d, $5.11 \times 10^{-4} \text{ kmol per kg Rh s}^{-1}$. This clearly indicates that the rate enhancement is due not to an intrinsic increase in activity of the exchanged complex $\text{HRh}(\text{CO})(\text{TPPTS})_{3-x}(\text{PPh}_3)_x$, but to promotion of interfacial catalysis.

The question of leaching of rhodium to the organic phase and its catalytic activity was also investigated. Recycling the organic phase had no activity for hydroformylation, and analysis of rhodium content in the organic phase was found to be $<5 \text{ p.p.m.}$ (by ultraviolet spectrophotometry) under the conditions used here.

We also tested the approach for biphasic hydroformylation of a water-soluble olefinic substrate (for example, allyl alcohol) using $\text{HRh}(\text{CO})(\text{PPh}_3)_3$ catalyst in the organic phase. Here, the effect of TPPTS as a promoter in the catalyst-immiscible aqueous phase was studied. The results (Fig. 2) indicate a five-fold increase in the rate of hydroformylation of allyl alcohol in the presence of TPPTS. Thus, our approach of 'promoter ligand' enhancement can also prove effective 'in reverse'. □

Received 14 July 1994; accepted 10 January 1995.

1. Parshall, G. W. *Homogeneous Catalysis* (Wiley Interscience, New York, 1980).
2. Kuntz, E. G. *CHEMTECH* **17**, 570–575 (1987).
3. Yu-Ren Chin *Oxo Alcohols* (Standard Res. Inst. Rep. 21C, SRI International, Menlo Park, California, 1986).
4. Hablot, I., Jenck, J., Casamatta, G. & Delmas, H. *Chem. Engng Sci.* **47**, 2689–2694 (1992).
5. Arhancet, J. P., Davis, M. E., Merola, J. S. & Hanson, B. E. *Nature* **339**, 454–455 (1989).
6. Larpent, C., Debard, R. & Patin, H. *Inorg. Chem.* **26**, 2922–2924 (1987).
7. Waller, F. J. *J. Molec. Catal.* **31**, 123–136 (1985).
8. Borowski, A. F., Cole-Hamilton, D. J. & Wilkinson, G. *Nouv. J. Chim.* **2**, 137–144 (1978).
9. Arhancet, J. P., Davis, M. E., Merola, J. S. & Hanson, B. E. *J. Catal.* **121**, 327–339 (1990).

ACKNOWLEDGEMENTS. We thank G. Samuel for assistance with NMR spectra. This work was supported by the Indo-French Center for Promotion of Advanced Research (Centre Franco-Indien pour la Promotion de la Recherche Avancée), New Delhi.

Modulation of hemispheric sea-ice cover by ENSO events

Per Gloersen

Oceans and Ice Branch, Laboratory for Hydrospheric Research, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

THE El Niño/Southern Oscillation (ENSO) is a quasiperiodic variation in climate which arises from a complex interaction between the tropical Pacific Ocean and the atmosphere¹. ENSO events, which occur every two to seven years, are the largest source of interannual variability of temperature and precipitation on a global scale, although their effects are most profound in the tropics¹. Observations of sea-ice margins have been used to monitor global climate changes on timescales of greater than a decade², and there is some evidence for interannual variations in records of sea-ice cover³. But short-term changes in sea-ice cover are masked by pronounced seasonal variations, making it difficult to correlate them with specific climate phenomena. Using a multiple-window harmonic analysis technique^{4–8}, I show here that time series of sea-ice cover from the Arctic and Antarctic contain statistically significant quasi-biennial and quasi-quadrennial periodicities that agree well with variations in the ENSO index. The response of sea ice to these two frequency components varies greatly for different regions.

Quasi-biennial and quasi-quadrennial periodicities attributed to the ENSO appear as significant spectral components in the time series of the length-of-day (LOD) parameter^{9–11}, as determined from very-long-base interferometer, space geodetic, satellite laser ranging, and Global Positioning Satellite (GPS) data¹². These spectral components also appear in the atmospheric angular momentum arising from zonally averaged global winds^{9,10}. With such clear global effects of ENSO, one might suspect that evidence of the ENSO should also clearly appear in global sea-ice coverage. Weak correlations have been reported earlier in the Bering Sea¹³, and the Indian and Pacific oceans¹⁴. But to date, a clear observation has eluded investigators because of the strong seasonal signal in sea-ice cover as well as localized weather events, which most certainly affect the nearby sea-ice coverage.

The sea-ice cover information discussed here was obtained from an analysis of microwave radiance data collected by the scanning multichannel microwave radiometer (SMMR) operating on board the NASA Nimbus 7 satellite from 26 October 1978 to 20 August 1987³.

A hint that there may be a Southern Oscillation index/sea ice correlation is shown in Fig. 1c, in which a residual obtained by removing the seasonal cycle from a time series of the Arctic sea-ice extents¹⁵ has been smoothed with a running annual average. For comparison, the SOI anomaly (obtained from the NOAA Climate Analysis Center) and the LOD parameter¹², both also smoothed with an annual running average, are shown in Fig. 1a and b, respectively. As noted earlier^{9,10}, the SOI signal during 1982–84 (and 1986–87) seems to lead the

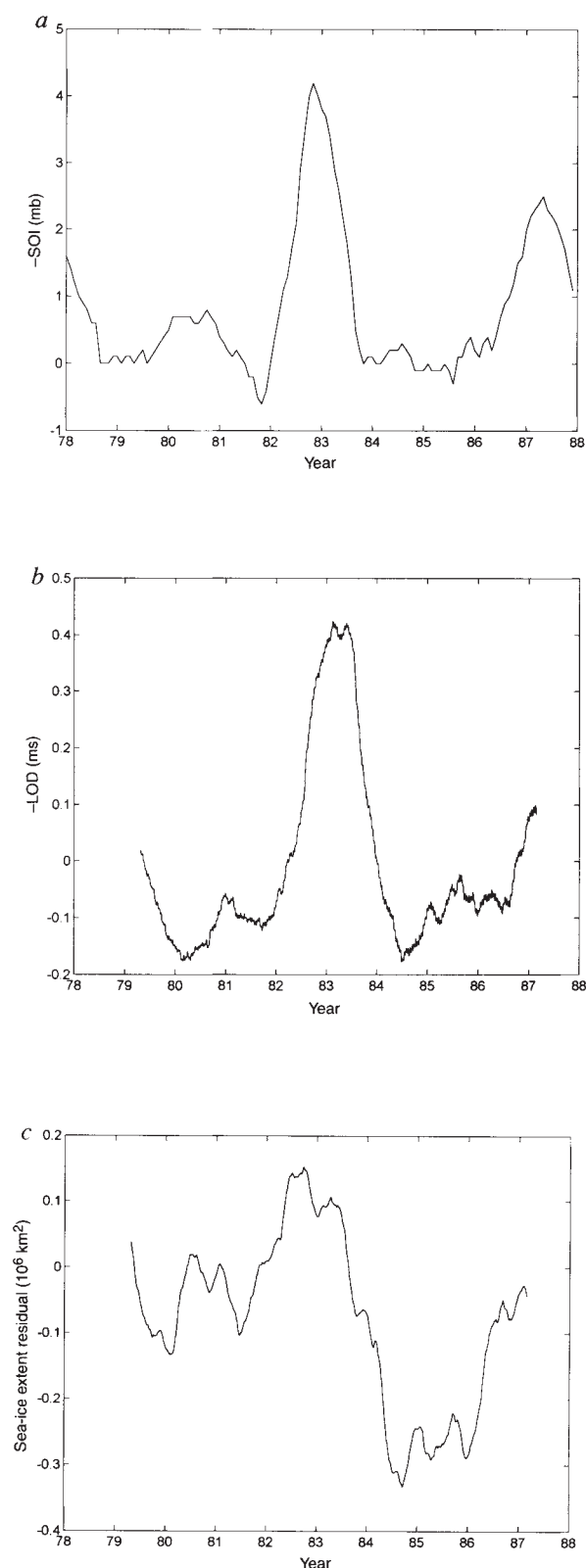


FIG. 1 Running annual averages of: a, the negative of the Southern Oscillation Index (SOI) obtained from the NOAA Climate Analysis Center; b, the length-of-day (LOD) parameter¹²; c, the Arctic sea-ice extent residual¹⁵ (obtained by removing the mean and modelled seasonal cycle from the record of sea-ice extents). The 1983 peak in the negative SOI corresponding to the El Niño event of that year leads that of the LOD by several months. The ice extent has a double peak in 1982–1983, with the second peak approximately in coincidence with the LOD. The first ice-extent peak in 1982 leads both the negative SOI and LOD peaks.

LOD by several months. In the Arctic sea-ice extent residual record, there is a corresponding double peak in the 1982–84 interval.

To clarify comparisons of this sort, I use here a technique developed by others for investigating seismic, atmospheric carbon dioxide, and other time series records^{4,8} to examine the unsmoothed time series of ice extents and ice areas (the ice extent less the amount of open water within the ice margin) in a nine-year record obtained from the SMMR³. I have also applied the technique to the LOD parameter, which has been suggested as a proxy for the ENSO events^{9,10}. This technique permits investigation within a narrow spectral range to exclude interfering events, with the result that the ENSO spectral components are more clearly seen than before in the LOD and for the first time in the sea-ice covers. This technique is summarized in Box 1.

When this procedure is applied to the time series of Arctic ice extents from SMMR^{3,15} with values of τ from 1 to 2,920 days, many spectral lines are obtained in which the annual cycle and four of its harmonics predominate. In Fig. 2a, I have plotted only the range 400–2,920 days (converted to 1–8 years on the abscissa) in order to emphasize sea-ice oscillations in the range of the ENSO components. A quasi-quadrennial (QQ) component is the strongest line in the spectral distribution of the F -statistic with a value of ~ 4.5 , well above the 95% confidence level at the F -value of 3.75. There are quasi-triennial (QT) and quasi-biennial (QB) peaks of lesser significance (an F -test value of 0.73 represents the 50% confidence level), and some even smaller, but distinct lines at periods of less than two years. The tail of the annual cycle line can be seen at the beginning of the distribution. The spectrum of the Arctic sea-ice area time series³ in Fig. 2a is very similar to that of the Arctic extent, but with generally smaller values of F , consistent with the idea that the marginal sea-ice zone might be more sensitive to ENSO forcing than the interior pack. The situation is similar in the Antarctic (Fig. 2b), but the F -values are generally smaller than in the Arctic, and the QT component is much more prominent.

To permit a direct comparison between SMMR and LOD¹² data, I have performed the above analysis on a subset of the LOD data corresponding in time to the SMMR data. The LOD power spectrum has been presented elsewhere^{9,10}; here, I show the spectral distribution of the F -test parameter of the LOD in Fig. 1a and b. There is a close match between the QB and QQ

BOX 1 Multiple-window harmonic analysis scheme

Following procedures described in detail elsewhere^{10–14,16}, eight different filtered data vectors are produced in the frequency domain by the Fourier transformations:

$$y_k(\tau) = \sum_{t=1}^{3,224} V_k(t)y(t) e^{-2\pi i t/\tau} \quad k=1, 2, \dots, 8 \quad (1)$$

where $V_k(t)$ is a data window of bandwidth $2W$ (with W chosen to be $4/N$, N being the number of observations) computed by taking a singular value decomposition of the sinc matrix⁴, and $1/\tau$ is the central frequency of the spectral band. y is the data vector. The sum over time t is taken on six-day intervals for the sea-ice data, to match the revisit time of the SMMR, and on one-day intervals for the LOD data. The coefficients of the model solution are obtained from the weighted sum of the eight differently-windowed and transformed data vectors, y_k :

$$a(\tau) = \sum_{k=1}^8 \frac{y_k(\tau)V_{0k}}{\sum_{k=1}^8 V_{0k}^2} \quad (2)$$

where V_{0k} is a normalizing factor obtained by summing the columns of $V_k(t)$.

The unexplained variance (in the frequency domain, and within the spectral bandwidth, $2W$) is then obtained by summing the squares of the differences between each of the eight windowed and transformed data vectors and the weighted coefficients of the model solution:

$$e^2(\tau) = \sum_{k=1}^8 \{ [y_k(\tau) - a(\tau)V_{0k}][y_k(\tau) - a(\tau)V_{0k}]^* \} \quad (3)$$

where $*$ indicates the complex conjugate.

To obtain a test of statistical significance for the spectral lines that result from the model solution amplitudes, $a(\tau)$, I have used an F -test statistic, which is written in terms of the ratio of the explained to unexplained variance within the bandwidth, $\pm W$:

$$F(\tau) = \{(n-1)[a(\tau)a^*(\tau)] \sum_{k=1}^8 V_{0k}^2\} / e^2(\tau) \quad (4)$$

In the present case, $n=8$, the number of independent samples of the data vector in the frequency domain. The random variable F has 2 and $2n-2$ degrees of freedom if the data are normally distributed.

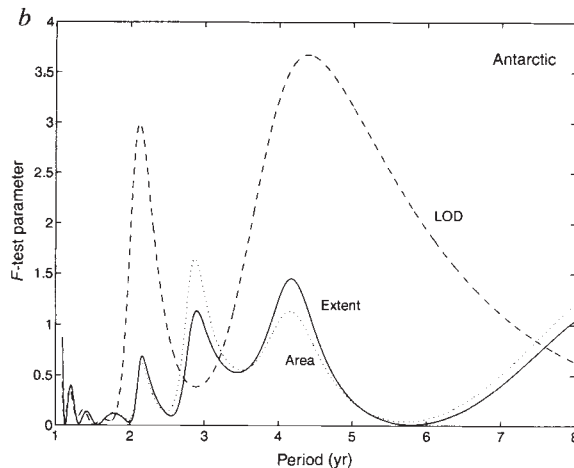
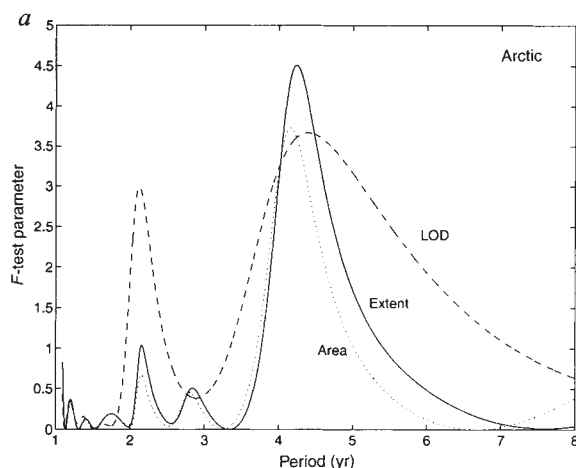


FIG. 2 Spectral distribution of the F -test periodicity in the nine-year time series of sea-ice area and extent in the Arctic (a) and the Antarctic (b). The LOD spectrum is superimposed for comparison. In the Arctic, the quasi-quadrennial (QQ) peak, at about 4.2 years, for the ice extent has a confidence level of 97%; those of the ice area and LOD, 94.5%. The strong quasi-biennial (QB) peak in the LOD shows up in the sea-ice area and extent with lower confidence levels, 60% or less. The

F -test values were obtained by first filtering the data with multiple-window tapers devised with prolate spheroid functions¹⁶ and then performing a discrete Fourier transform on the filtered data to obtain the complex coefficients of the model fit to the data. The F -test is the ratio of the explained variance (modelled data) to the unexplained variance (observed data minus the modelled data) within the bandwidth, $\pm W$, $\sim 1/806 \text{ d}^{-1}$.

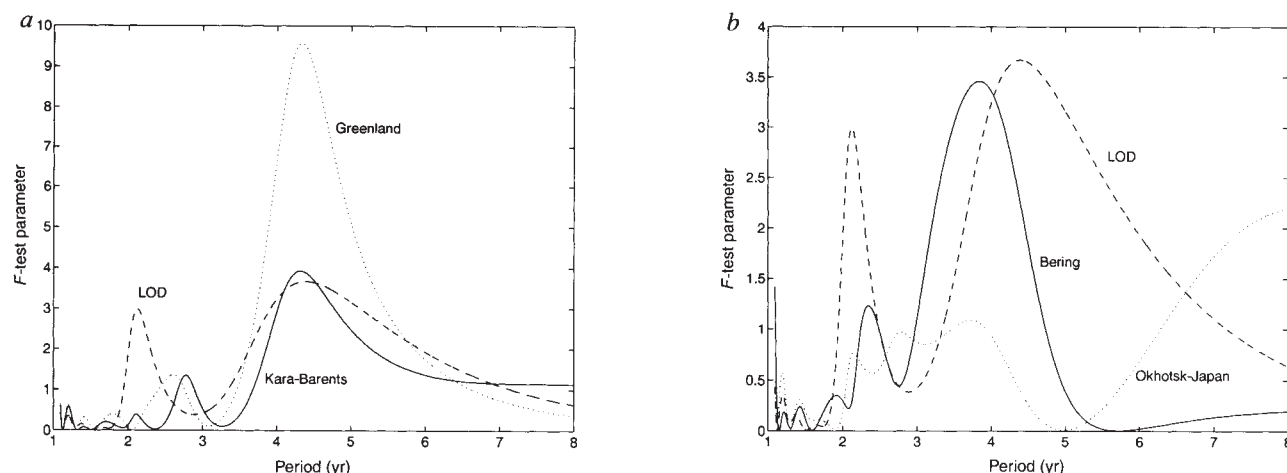


FIG. 3 Same as Fig. 2, except for the sea-ice area only in: a, the Greenland and Kara-Barents regions; b, the Bering and Okhotsk-Japan regions³. The confidence level of the QQ component in the Kara-Barents

region is ~95%; in the Greenland Sea, well above 99%. The QB and QQ components in the Bering and Okhotsk-Japan regions are shifted to higher and lower periods, respectively, than those of the LOD.

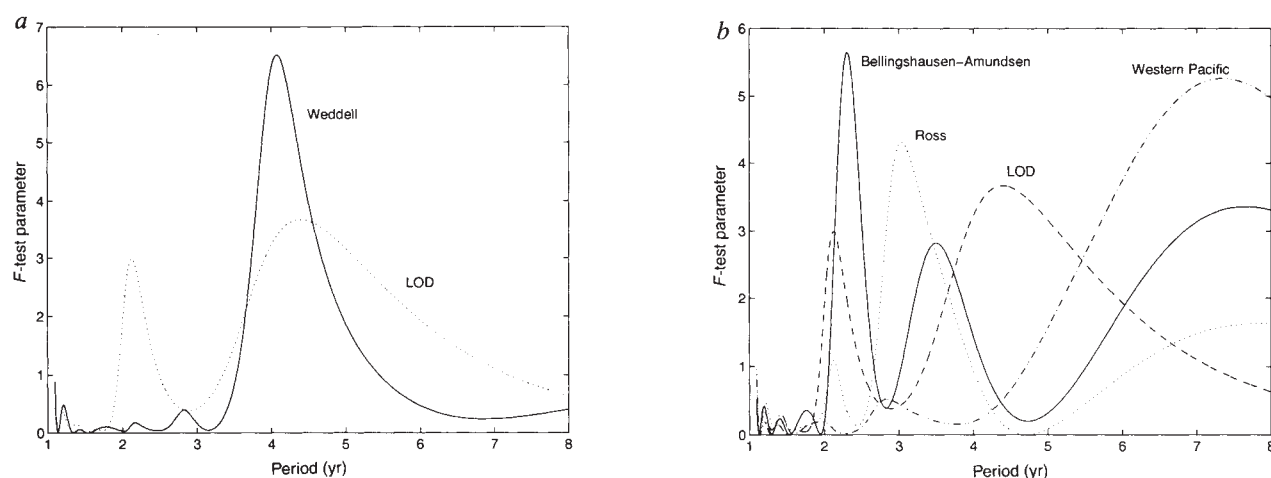


FIG. 4 Same as Fig. 2, except for the sea-ice area only; a, in the Weddell Sea sector³; b, in the Bellingshausen-Amundsen, Ross and Western Pacific sectors³. In the Weddell Sea sector, the QQ component predominates, whereas in the Bellingshausen-Amundsen Seas sector, the QB

component is dominant. There is a strong quasi-triennial (QT) component in both the Bellingshausen-Amundsen and Ross Sea sectors. There are no substantial QB, QT, or QQ components in the Western Pacific sector, nor in the Indian Ocean sector³ (not shown).

components of the LOD and those of the Arctic and the Antarctic. The QT component in the Arctic and Antarctic is missing in the LOD spectrum because the LOD time series used here is on exactly one-day intervals. The QT component appears to be an alias frequency resulting from the beating of the SMMR revisit frequency against the S_1 (and/or S_2) oceanic/atmospheric tidal frequency.

Pursuing the idea of greater ENSO sensitivity in the marginal sea-ice zones, I have selected for similar analysis the Arctic regions, defined elsewhere³, that include some of the perimeter seas. The most prominent indication of the QQ component, about twice the F -value of that in the Arctic ice extent (Fig. 2a), occurs in the time series of the Greenland Sea ice areas (Fig. 3a), where also the QB and QT components are not as distinct. The QQ component in the Kara-Barents region (Fig. 3a) is comparable to that of the Arctic extent, and the QT component is more distinct. In the Bering region (Fig. 3b), the QQ and QT components appear to have shorter periods than the Arctic sum

by about half a year. In the Okhotsk-Japan region, the QQ, QT and QB components run together and are all below the 70% confidence level (Fig. 3b).

Turning to the Antarctic, the Weddell Sea sector³ shows the QQ component as the only significant peak (Fig. 4a). The QB component is the strongest in the Bellingshausen-Amundsen sea sector³ (Fig. 4b). The Ross Sea sector³ shows a strong QT peak and a weaker QB peak.

It is clear that variations in the areal coverage of sea ice on global, hemispheric and regional bases contain periodicities of quasi-biennial and quasi-quadrennial components, very similar to the variations in the length-of-day, El Niño-Southern Oscillation index and atmospheric angular momentum parameters^{9,10}. The average response to these two components is more significant in the Arctic than in the Antarctic. However, the regional responses to these two components vary greatly. The QQ oscillation is most significant in the Greenland Sea region and almost as significant in the Weddell Sea, whereas the QB com-

ponent is most significant in the Bellingshausen–Amundsen Sea sector. □

Received 11 August; accepted 8 December 1994.

1. Philander, S. G. H. *El Niño, La Niña and the Southern Oscillation* (Academic, New York, 1990).
2. Lamb, H. H. *Climate: Past, Present, and Future* (Methuen, London, 1972).
3. Gloersen, P., Campbell, W. J., Cavalieri, D. J., Comiso, J. C., Parkinson, C. L. & Zwally, H. J. *Arctic and Antarctic Sea Ice, 1978–1987: Satellite Passive-Microwave Observations and Analysis* (SP-511, NASA, Washington DC, 1992).
4. Lindberg, C. R. thesis, Univ. California, San Diego (1986).
5. Park, J., Lindberg, C. R. & Vernon, F. L. *J. geophys. Res.* **92**, 12675–12684 (1987).
6. Park, J., Lindberg, C. R. & Thomson, D. J. *Geophys. J. R. astr. Soc.* **91**, 755–794 (1987).
7. Lindberg, C. R. & Park, J. *Geophys. J. R. astr. Soc.* **91**, 795–836 (1987).
8. Kuo, C., Lindberg, C. R. & Thomson, D. J. *Nature* **343**, 709–714 (1990).
9. Dickey, J. O., Marcus, S. L. & Hide, R. *Nature* **357**, 484–488 (1992).
10. Dickey, J. O., Marcus, S. L., Eubanks, T. M. & Hide, R. *Interactions Between Global Climate Systems, The Legacy of Hann 141–155* (Geophys. Monogr. 75, Am. Geophys. Union, Washington DC, 1993).
11. Chao, B. F. *Science* **243**, 923–924 (1989).
12. Gross, R. S. & Stepe, J. A. *A Combination of Earth Orientation Data: SPACE93* (Observatoire de Paris, 1993).
13. Niebauer, H. J. *J. geophys. Res.* **93**, 5051–5068 (1988).
14. Simmonds, I. & Jacka, T. H. *J. Clim.* (in the press).
15. Gloersen, P. & Campbell, W. J. *Nature* **352**, 33–36 (1991).
16. Thomson, D. J. *Proc. IEEE* **70**, 1055–1095 (1982).

ACKNOWLEDGEMENTS. I thank E. Mollo-Christensen, D. J. Cavalieri, S. Hakkinen and C. J. Kobinsky for helpful suggestions. This work was sponsored by the Science Division of the Office of Mission to Planet Earth at NASA Headquarters.

Climate-related variations in denitrification in the Arabian Sea from sediment $^{15}\text{N}/^{14}\text{N}$ ratios

Mark A. Altabet*, Roger Francois*, David W. Murray† & Warren L. Prell†

* Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

† Department of Geological Sciences, Brown University, Box 1846, Providence, Rhode Island 02912, USA

DENITRIFICATION—the process by which nitrate is reduced to gaseous nitrogen species (usually N_2 or N_2O)—is the dominant mechanism for removal of fixed nitrogen from the biosphere. In the oceans, denitrification is mediated by bacteria in suboxic environments and, by controlling the supply of fixed nitrogen, is an important limiting factor for marine productivity^{1–3}. Denitrification produces substantial ^{15}N enrichment in subsurface nitrate^{4–6}, which is reflected in the isotopic composition of sinking particulate nitrogen⁷; sediment $^{15}\text{N}/^{14}\text{N}$ ratios in regions with suboxic water columns may therefore provide a record of past changes in denitrification intensity. Here we report nitrogen isotope data for sediment cores from three sites in the Arabian Sea. At all three sites we find large, near-synchronous downcore variations in $^{15}\text{N}/^{14}\text{N}$, which are best explained by regional changes in the isotopic composition of subsurface nitrate, and hence denitrification. Moreover, these variations are synchronous with Milankovitch cycles, thereby establishing a link with climate. We argue that these large, climate-linked variations, in a region that contributes significantly to global marine denitrification, are likely to have perturbed marine biogeochemical cycles during the Late Quaternary period.

The $^{15}\text{N}/^{14}\text{N}$ ratio of particulate organic matter sinking out of the surface ocean is principally a function of the $\delta^{15}\text{N}$ (in ‰ relative to atmospheric N_2 ; see Fig. 2 legend) of source nitrate (NO_3^- , transported vertically from subsurface waters) and isotopic fractionation occurring during partial utilization of this nutrient by phytoplankton^{8,9}. Recent results have shown this surface-generated isotopic signal to be subsequently recorded in deep-sea sediment¹⁰. In the equatorial Pacific and the Southern Ocean, large latitudinal gradients in $\delta^{15}\text{N}$ for both near-surface

ocean particulate organic matter and core-top sediments are inversely correlated with surface $[\text{NO}_3^-]$. In these regions, variations in $\delta^{15}\text{N}$ are driven by variations in nutrient utilization because the $\delta^{15}\text{N}$ of subsurface NO_3^- is fairly constant throughout most of the ocean⁶. But in oceanic regions with suboxic subsurface waters (that is, strong oxygen-minimum zones), denitrification also produces substantial enrichment in ^{15}N for the residual NO_3^- (refs 5, 6) which extends regionally⁶ as a result of horizontal advection and mixing. By producing regional increases in the $\delta^{15}\text{N}$ of subsurface NO_3^- , water-column denitrification strongly influences the $\delta^{15}\text{N}$ of sinking particulate organic matter and sediments in these domains^{7,11}.

We have examined sediment cores at three sites in the Arabian Sea to isolate past variations in these effects (Fig. 1). Summer monsoon winds produce upwelling along the Oman coast out to ~350 km of shore. At this time when productivity and downward particle fluxes are high^{12,13}, the corresponding northwest to southeast gradient in surface nutrients¹⁴ may be expected to produce spatial variations in the $\delta^{15}\text{N}$ of near-surface particulate organic matter. Near-core-top sediments do show a 1‰ increase from the margin outwards to the Owen ridge (350 km offshore; Fig. 2). The Owen ridge (RC27-61) and nearshore Oman margin sites (ODP site 723, hole B and RC27-24) are at the edge and well within the upwelling system, respectively. Past changes in the gradient in $\delta^{15}\text{N}$ between them could reflect changes in the

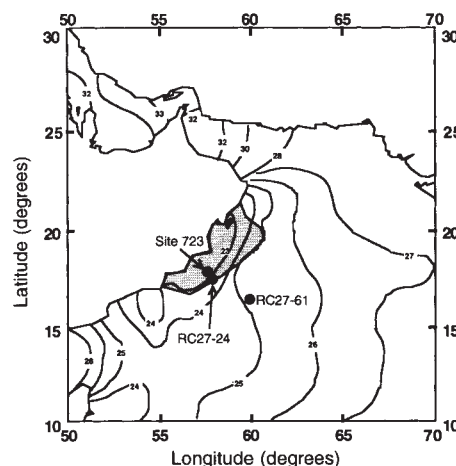


FIG. 1 Core locations in the northwest Arabian Sea. RC27-24 (17° 43.1' N, 57° 49.2' E; 1,416 m depth) and ODP site 723 (18° 3.1' N, 57° 36.6' E; 808 m depth) are from the Oman margin. RC27-61 (16° 39.5' N, 59° 31.4' E) is located 350 km offshore at 1,893 m depth. Contours of summer sea surface temperature show the gradient in monsoon-driven upwelling intensity¹⁶. The shaded area shows the region of high-nutrient surface waters ($[\text{PO}_4^{3-}] > 1 \mu\text{M}$) associated with upwelling¹⁴. Suboxic water between 100 and 1,000 m depth is present throughout most of this region, including over our core sites. Comparing results for the two cores in the Oman margin shows that depositional environment has little effect on sedimentary $\delta^{15}\text{N}$ in this region. Site 723 at 808 m depth is overlain by the oxygen-minimum zone; core RC27-24, 24 km to the southeast at 1,416 m water depth, is overlain by oxic water. Despite large differences between site 723 and RC27-24 in particulate nitrogen content (for example, 1% and 0.45%, respectively, at core top; the sedimentary nitrogen accumulation rates show similar differences), the downcore $\delta^{15}\text{N}$ variations are very similar for these two sites (Fig. 2). For the Late Holocene, values are indistinguishable. Considering that these sites are sufficiently close to receive the same input from near-surface waters, the differences in %N and N accumulation rate are probably related to contrasts in nitrogen preservation and/or along-bottom transport or trapping of sediment. In either case, these results indicate that possible past changes in depositional environment are unlikely to have significantly influenced the $\delta^{15}\text{N}$ recorded in this region.

FIG. 2 Variation in sedimentary $\delta^{15}\text{N}$ (solid lines) over the past 150 kyr for the two Oman margin cores (ODP site 723, hole B and RC27-24; top panel) and over the past 430 kyr for the Owen ridge core (RC27-61; bottom panel). For comparison, the planktonic foraminiferal $\delta^{18}\text{O}$ record (dashed lines) for each core is shown (from *Pulleniatina obliquiloculata* for site 723³⁷, from *Globigerinoides sacculifer* for the other two). Glacial stages are indicated by shaded regions^{15,19}. The age model for RC27-61 is as previously described³⁸; the age models for RC27-24 and site 723 have been slightly improved from those previously published²⁰ based on carbonate stratigraphy (D.W.M. and W.L.P., unpublished data). The youngest sample measured in piston core RC27-61 is 3.7 kyr corresponding to a depth of 2 cm. $\delta^{15}\text{N}$ values were measured on bulk sediment and expressed relative to atmospheric N_2 with a precision of $\pm 0.2\text{‰}$; $\delta^{15}\text{N} = [({}^{15}\text{N}/{}^{14}\text{N})_{\text{sample}}/({}^{15}\text{N}/{}^{14}\text{N})_{\text{atmosphere}} - 1] \times 1,000$. Particulate nitrogen was automatically converted to N_2 using a CN analyser which was passed to an isotope-ratio mass spectrometer (Finnigan MAT 251) in a continuous flow of He carrier. Variations in sedimentary $\delta^{15}\text{N}$ reflect changes in the $\delta^{15}\text{N}$ of particulate nitrogen deposited from overlying water. Assuming short residence time for particulate nitrogen in surface waters its $\delta^{15}\text{N}$ is given by:

$$\delta^{15}\text{N}_{\text{Sinking POM}} = \delta^{15}\text{NO}_3^-_{\text{subsurf.}} - (\epsilon_u \ln(1-u)) - \epsilon_u$$

where ϵ_u is the fractionation factor associated with phytoplankton uptake, POM is particulate organic matter and u is the relative utilization of NO_3^- ($1 - [\text{NO}_3^-]_{\text{obs}}/[\text{NO}_3^-]_{\text{initial}}$; a function of physical supply and biological removal). This equation is the combination of the Rayleigh fractionation equation:

$$\delta^{15}\text{NO}_3^-_{\text{surf.}} = \delta^{15}\text{NO}_3^-_{\text{subsurf.}} - \epsilon_u \ln(1-u)$$

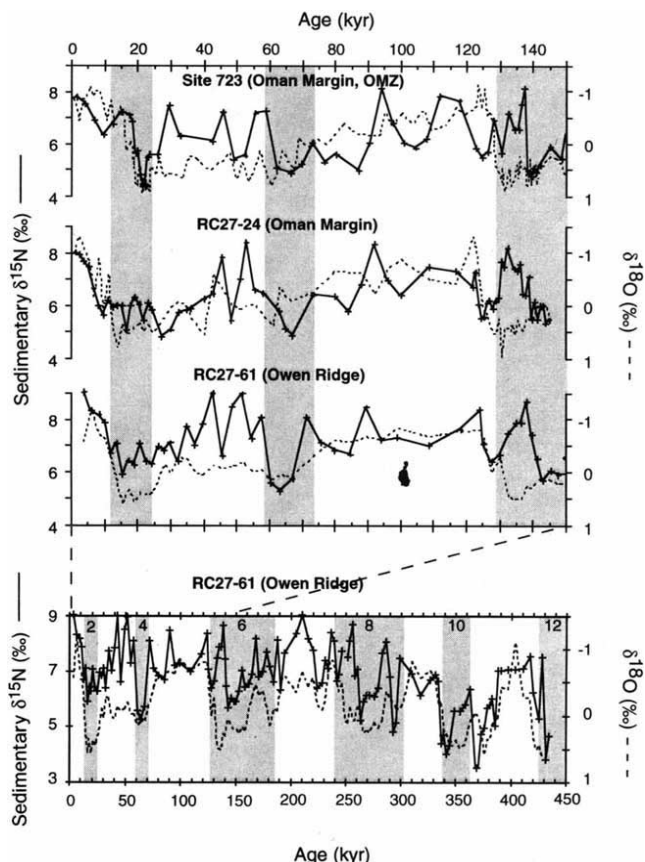
and the instantaneous product equation:

$$\delta^{15}\text{N}_{\text{Sinking POM}} = \delta^{15}\text{NO}_3^-_{\text{surf.}} - \epsilon_u$$

spatial pattern of NO_3^- utilization. Changes in $\delta^{15}\text{N}$ occurring in concert, on the other hand, would reflect changes in the $\delta^{15}\text{N}$ of subsurface NO_3^- and hence denitrification intensity.

Over the last 150 kyr, the Owen ridge core (RC27-61) generally has $\delta^{15}\text{N}$ values higher than the Oman margin cores; this could be interpreted as reflecting increased surface nutrient utilization with distance from the upwelling centre. Past changes in this offshore gradient in $\delta^{15}\text{N}$ may reflect variations in the upwelled supply of NO_3^- . The predominant $\delta^{15}\text{N}$ events observed, however, occur in near-synchrony in all three cores (the small temporal offsets between maxima and minima are probably a consequence of discrete sample intervals and age-model uncertainties). They are clearly regional in nature; we interpret these events as reflecting past changes in the $\delta^{15}\text{N}$ of subsurface NO_3^- and thus denitrification. Inspection of the common downcore variations in sedimentary $\delta^{15}\text{N}$ indicates linkages to climate cycles. At core top, $\delta^{15}\text{N}$ values are near maximal, decreasing to minimal values ~ 20 kyr ago (isotope stage 2) and are evidence for greatly reduced denitrification during the Last Glacial Maximum. At 65 kyr (isotope stage 4), a less-intense maximum in ice volume corresponds to another $\delta^{15}\text{N}$ minimum. The previous deglaciation (~ 130 kyr) is also marked by rising $\delta^{15}\text{N}$ values, but a $\delta^{15}\text{N}$ peak occurs at ~ 135 kyr that is not matched in the planktonic foraminiferal $\delta^{18}\text{O}$ record (which reflects both global ice volume and sea surface temperature).

In the longer RC27-61 $\delta^{15}\text{N}$ record (to 430 kyr, Fig. 2), glacial stages also usually contain minima in $\delta^{15}\text{N}$, although significant excursions occur at periods shorter than the 100-kyr cycle that dominates ice volume. Spectral analysis shows significant peaks in spectral density associated with the variation in Earth's orbital parameters; at 23 and 41 kyr as well as 100 kyr (Fig. 3). ETP (eccentricity, tilt, precession)¹⁵ is an index of variations in Earth's orbital parameters that control the temporal distribution of Northern Hemisphere solar insolation. More intense summer monsoons have been linked to increased Northern Hemisphere insolation but with a significant lag^{16,17}. Cross-spectra with the $\delta^{15}\text{N}$ record demonstrate significant coherency at the 100-, 41-



and 23-kyr periods (Fig. 3) with $\delta^{15}\text{N}$ significantly lagging at each period by 6–9 kyr.

Proximal forcings for denitrification intensity can be explored using cross-spectral analysis with other palaeorecords. The relative abundance of the planktonic foraminifera *Globigerina bulloides* is a specific index of upwelling intensity¹⁸ and cross-spectra with $\delta^{15}\text{N}$ show highly significant coherency at the 41- and 23-kyr periods. The lack of a strong 100-kyr period for the percentage of *G. bulloides* in RC27-61 has been attributed to the effects of dissolution¹⁹. Focusing on the 23-kyr period, the RC27-61 $\delta^{15}\text{N}$, RC27-61 %*G. bulloides* and site 723 %*G. bulloides* records are nearly in phase, suggesting a strong regional link between upwelling and denitrification (Fig. 3 legend). These records are also nearly in phase with variations in opal flux and foraminifera fragmentation; all these records are considered productivity indices (Fig. 3 legend)^{16,20}. These results are consistent with denitrification responding rapidly to changes in monsoon-induced upwelling intensity which varies most strongly at the precessional period (23 kyr)¹⁶.

The $\delta^{15}\text{N}$ and associated data indicate that the Arabian Sea has exhibited large, climate-driven variations in denitrification probably forced directly by changes in intensity of the summer monsoon and associated upwelling. Variation in the extent and intensity of the oxygen-minimum zone is a likely mechanistic link between upwelling and denitrification. Upwelling stimulates high levels of productivity and particle flux in the Arabian Sea¹², which is a major contributor to the maintenance of this zone²¹. But because warm, low-oxygen source waters to the oxygen-minimum zone from the Persian Gulf and Gulf of Aden are also important for formation of suboxic water²¹, past changes in their source strength may have also contributed to variations in denitrification. For example, the Persian Gulf is known to have dried up completely during the last glaciation as a result of lowering of sea level below sill depth²². Thus part of the variation in denitrification may be forced by sea level and hence be more directly related to ice volume, accounting for strong 100-kyr periodicity in $\delta^{15}\text{N}$ (Fig. 3).

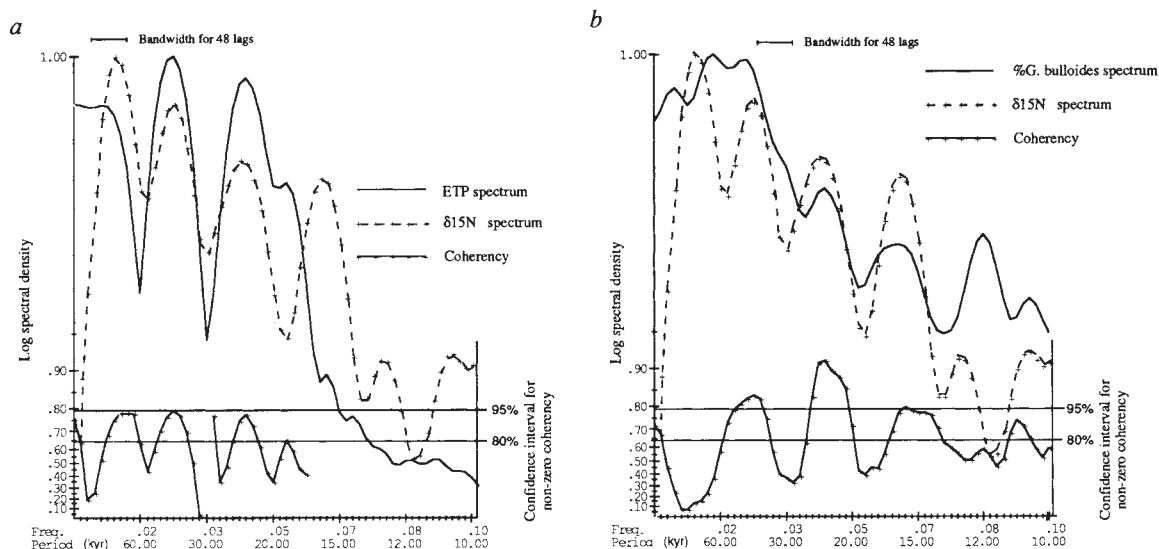


FIG. 3 Spectral analysis of $\delta^{15}\text{N}$ data for core RC27-61 (a 430-kyr record) showing major periods of variation at orbital periods of 100, 41 and 23 kyr. Spectral density (σ^2/f , where σ^2 = variance and frequency $f=1/\text{period}$) is normalized and plotted on a logarithmic scale. Bandwidth is $0.009257 \text{ (kyr}^{-1}\text{)}$ and interpolated time interval (Δt) is 3 kyr. Cross-spectra are used to show coherence with ETP (time series of summed Earth's eccentricity, tilt and precessional parameters, a) and the RC27-61 %G. *bulloides* record (indicator of upwelling intensity, b). The latter data are from ref. 16. Coherence is plotted on a hyperbolic arctangent scale and the 80% and 95% confidence intervals are shown (0.644 and 0.795, respectively). Before analysis, the linear trend in $\delta^{15}\text{N}$ between 250 and 430 kyr (Fig. 2) was removed to identify better

the 100-kyr periodicity. The $\delta^{15}\text{N}$ record lags ETP by 6.8 ± 6.0 , 5.5 ± 2.1 and 9.0 ± 1.3 kyr at the 100-, 41- and 23-kyr periods, respectively. At the 23-kyr period, the $\delta^{15}\text{N}$ record shows little or no lag with upwelling indicators, namely the RC27-61 %G. *bulloides* record (1.3 ± 1.4 kyr) and ODP site 723 %G. *bulloides* record (3.0 ± 1.7 kyr)²⁰. Similarly there is little lag at the 23-kyr period with productivity indices, namely the RC27-61 foraminifera fragments record (0.6 ± 1.0 kyr)³⁹, and ODP site 722 (near RC27-61 in the Owen ridge) opal flux (1.8 ± 1.3 kyr)¹⁶. We have greater confidence in lags estimated between records from the same core, because age model uncertainties do not influence these calculations.

In regions other than those influenced by denitrification, sub-surface NO_3^- has $\delta^{15}\text{N}$ values⁶ of $\sim 5\text{--}6\text{‰}$ which are also reflected³ in surface sediment $\delta^{15}\text{N}$. As sedimentary $\delta^{15}\text{N}$ in the Arabian Sea drops to values as low as this, a complete collapse of denitrification is suggested, for example, during parts of stages 2, 4, 6, 8, 10 and 12 (Fig. 2). Through exchange with other ocean basins, this decrease in $\delta^{15}\text{N}$ should also produce a comparatively smaller ocean-wide change in average $\delta^{15}\text{N}$ for NO_3^- , which may have been detected previously³. Because the Arabian Sea accounts for a large proportion of marine denitrification, its cessation or near-cessation would have had ocean-wide consequences. Water-column denitrification accounts for about half of the total marine denitrification ($\sim 120 \text{ Tg yr}^{-1}$)¹, the balance occurring in shelf sediments^{2,23}. Denitrification in the Arabian Sea has been reported as being between 12 (ref. 40) and 30 (ref. 24) Tg yr^{-1} and thus could account for up to 10–25% of total marine denitrification (the difference is due to varying estimates for the residence time of water in the suboxic zone). The significance of such a reduction for fixed-nitrogen inventories depends on its duration relative to the residence time of fixed nitrogen in the ocean. Given recent nitrogen budgets¹, the oceanic residence time for fixed N is estimated at 5 kyr. Periods of minimum $\delta^{15}\text{N}$ are of this length or longer, allowing sufficient time for oceanic nitrogen to approach a new equilibrium with 10–25% higher concentrations, assuming all other sources and sinks are invariant. Depending on interactions with other limiting factors, the productivity of nutrient-depleted surface ocean waters may have increased, which in turn would lower²⁰ atmospheric P_{CO_2} . Changes in both water-column and benthic denitrification have been previously hypothesized as significant influences on marine biogeochemical cycles on glacial interglacial timescales^{3,26,28}. This work presents the first clear evidence from the sedimentary record for productivity-induced denitrification as a mechanism linking the ocean's nitrogen and climate cycles. Climate

responses by other water-column denitrification regions are currently under investigation²⁹.

The climate-linked changes in the extent and intensity of suboxic water in the Arabian Sea suggested by the $\delta^{15}\text{N}$ record have additional biogeochemical implications. For example, production of N_2O , a radiatively important gas that also influences stratospheric ozone concentration, is enhanced at low oxygen concentrations^{30,31}, such that the Arabian Sea is an important marine source of N_2O (0.3 Tg yr^{-1})^{32,33}. The minimum in atmospheric N_2O concentration observed for the Last Glacial Maximum³⁴ has been suggested to result from a reduction in terrestrial sources. We propose that reduction in the marine source of N_2O from the Arabian Sea may have contributed to reduced glacial N_2O . As N_2O from suboxic water is very isotopically enriched (T. Yoshinari and M.A.A., manuscript in preparation) compared to terrestrial sources³⁵, this effect may be detected in observations of the isotopic composition of ice-core N_2O given sufficient instrument sensitivity. □

Received 18 July; accepted 23 December 1994.

1. Codispoti, L. & Christensen, J. P. *Mar. Chem.* **16**, 277–300 (1985).
2. Devol, A. H. *Nature* **349**, 319–321 (1991).
3. Altabet, M. A. & Curry, W. B. *Global biogeochem. Cycles* **3**, 107–119 (1989).
4. Miyake, Y. & Wada, E. *Rec. oceanog. Works Japan* **11**, 1–6 (1971).
5. Cline, J. D. & Kaplan, I. R. *Mar. Chem.* **3**, 271–299 (1975).
6. Liu, K. K. & Kaplan, I. R. *Limnol. Oceanogr.* **34**, 820–830 (1989).
7. Schafer, P. & Ittekkot, V. *Naturwissenschaften* **80**, 511–513 (1993).
8. Altabet, M. A. *Deep-Sea Res.* **35**, 535–554 (1988).
9. Altabet, M. A. & Francois, R. in *Carbon Cycling in the Glacial Ocean: Constraints on the Ocean's Role in Global Change* (eds Zahn, R., Kaminski, M., Labeyrie, L. & Pederson, T. F.) 281–306 (Springer, Berlin, 1994).
10. Altabet, M. A. & Francois, R. *Global biogeochem. Cycles* **8**, 103–116 (1994).
11. Libes, S. M. & Deuser, W. G. *Deep-Sea Res.* **35**, 517–533 (1988).
12. Nair, R. R. et al. *Nature* **338**, 749–751 (1989).
13. Hacke, B. et al. *Deep-Sea Res.* **40**, 1323–1344 (1993).
14. Mortin, R. *Oceanographic Atlas of the International Indian Ocean Expedition* (National Science Foundation, Washington DC, 1971).

15. Imbrie, J. et al. in *Milankovitch and Climate Part 1* (eds Berger, A. L., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B.) 269–305 (Reidel, Hingham, 1984).
16. Clemens, S., Prell, W., Murray, D., Shimmield, G. & Weedon, G. *Nature* **353**, 720–725 (1991).
17. Prell, W. L. in *Milankovitch and Climate Part 1* (eds Berger, A. L., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B.) 349–366 (Reidel, Hingham, 1984).
18. Prell, W. L. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) 48–57 (Am. Geophys. Union, Washington DC, 1984).
19. Prell, W. L., Murray, D. W., Clemens, S. C. & Anderson, D. M. in *Synthesis of Results from Scientific Drilling in the Indian Ocean* (eds Duncan, R. A., Rea, D. K., Kidd, R. B., Rad, U. V. & Weissel, J. K.) 447–469 (Am. Geophys. Union, Washington, DC, 1992).
20. Anderson, D. M. & Prell, W. L. *Paleoceanography* **8**, 193–208 (1993).
21. Olson, D. B., Hitchcock, G. L., Fine, R. A. & Warren, B. A. *Deep-Sea Res.* **40**, 673–685 (1993).
22. Sarnthein, M. *Mar. Geol.* **12**, 245–266 (1972).
23. Christensen, J. P. W., Smethie, J. & Devol, A. H. *Deep-Sea Res.* **34**, 1027–1047 (1987).
24. Naqvi, S. W. A. *J. mar. Res.* **45**, 1049–1072 (1987).
25. Broecker, W. S. *Prog. Oceanogr.* **11**, 151–197 (1982).
26. Berger, W. H. & Keir, R. S. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 337–351 (Am. Geophys. Union, Washington DC, 1984).
27. Christensen, J. P., Murray, J. W., Devol, A. H. & Codispoti, L. A. *Globl biogeochem. Cycles* **1**, 97–116 (1987).
28. Codispoti, L. A. in *Productivity of the Ocean: Present and Past* (eds Berger, W. H., Smetacek, V. S. & Wefer, G.) 377–394 (Wiley, Chichester, 1989).
29. Ganeshram, R., Pedersen, T. F., Calvert, S. E. & Murray, J. W. *Mineralog. Mag.* **58A**, 313–314 (1994).
30. Goreau, T. J. et al. *Appl. envir. Microbiol.* **40**, 526–532 (1980).
31. Lipschultz, F. et al. *Nature* **294**, 641–643 (1981).
32. Naqvi, S. W. A. & Noronha, R. J. *Deep-Sea Res.* **38**, 871–890 (1991).
33. Codispoti, L. A. et al. in *Oceanography of the Indian Ocean* (ed. Desai, B. N.) 271–284 (Oxford & IBH, New Delhi, 1992).
34. Leunberger, M. & Siegenthaler, U. *Nature* **360**, 449–451 (1992).
35. Kim, K. & Craig, H. *Science* **262**, 1855–1857 (1993).
36. Hastenrath, S. & Lamb, P. J. *Climate Atlas of the Indian Ocean* (Univ. Wisconsin, Madison, 1979).
37. Niitsuma, N., Oba, T. & Okada, M. *Proc. ODP Sci. Res.* 321–335 (1991).
38. Clemens, S. & Prell, W. L. *Paleoceanography* **5**, 109–145 (1990).
39. Murray, D. W. & Prell, W. L. in *Unwelling Systems: Evolution since the Early Miocene* (eds Summerhayes, C. P., Prell, W. L. & Emeis, K. C.) 301–321 (Geological Soc., London, 1992).
40. Mantoura, R. F. C. et al. *Deep-Sea Res. II* **40**, 651–671 (1993).

ACKNOWLEDGEMENTS. We thank J. Donaghue for technical assistance and A. McKay for work on age-model improvements for ODP site 723 (hole B). The comments of L. Codispoti and J. Montoya considerably improved the manuscript. This work was supported by the US NSF.

A Plio-Pleistocene hominid from Dmanisi, East Georgia, Caucasus

L. Gabunia & A. Vekua

Institute of Paleobiology, Georgian Academy of Sciences,
9 Sulkhan-Saba, Tbilisi 380007, Georgia

ARCHAEOLOGICAL excavations at the mediaeval site of Dmanisi (East Georgia) revealed that the town was built on a series of deposits yielding Late Villafranchian mammalian fossils and led to the discovery in late 1991 of a well preserved early human mandible. Dmanisi, where excavations are being carried out by a joint expedition of the Archaeological Research Centre of the Georgian Academy of Sciences and the Römisch-Germanisches Zentralmuseum (Mainz, Germany), is located southwest of Tbilisi, at about 44° 20' N, 41° 20' E (Fig. 1). The fossils date to the latest Pliocene (or perhaps to the earliest Pleistocene), probably between 1.8 and 1.6 million years ago (Myr). Here we identify the mandible as belonging to the species *Homo erectus*, of which it is the earliest known representative in western Eurasia. It shows a number of similarities to the African and Chinese representatives of this species.

The geology of the Dmanisi region and the stratigraphy of the continental deposits yielding the human jaw and the associated vertebrate fauna have been described in preliminary reports^{1,2}. A basalt layer 80 m thick is overlain by about 4 m of alluvial (subaerial) deposits containing mammalian fossils and marked in places by soil horizons (Fig. 1b). The presence of *Kowalskia* sp., *Miomys* sp. (*M. reidi-pitomyoides* group)³, *Hypolagus brachygnathus* Kormos, *Pachyrocute* cf *perrieri* (Croizet et Jobert), *Archidiskodon meridionalis* Nesti, *Equus* cf *stenonis* Cocchi, *Dicerorhinus etruscus etruscus* (Falconer), and *Cervus perrieri*

Croizet et Jobert aligns the fauna with the Late Villafranchian of Europe, perhaps especially its earlier part⁴. A potassium-argon date of 1.8 ± 0.1 Myr was obtained on the basalt¹, which is in broad agreement with this age. Moreover, both the basalt and the bone-bearing layer show normal geomagnetic polarity^{1,5}, which may indicate that they belong to the end of the Olduvai Subchron (1.95–1.77 Myr), although more detailed studies would clarify this point. Numerous stone tools of 'Oldowan' appearance (Fig. 1c) also occur in the bone-bearing layers, at and above the level with the human mandible.

The Dmanisi hominid mandible consists of an almost entirely preserved corpus with complete and little-worn dentition (Fig. 2). The base is damaged and both rami are broken away. The general appearance is of a jaw with small teeth but a thick corpus, especially below the first and second molars (see Table 1). It is not possible to determine its sex unambiguously.

On the basis of the features detailed in Fig. 2, especially the overall size and robustness of the corpus and the teeth, shape of the symphysis, and dental proportions, the Dmanisi jaw represents an early form of *Homo*. Comparisons can be made with: east Asian populations of *H. erectus*, from Sangiran (Java) and Zhoukoudian (China); early African fossils assigned to that species or to *Homo ergaster* (mainly ER 730 and 992, WT 15000); later African *H. erectus* (OH 22, Tighenif (ex-Ternifine), Thomas Quarry 1); members of *Homo habilis* or *Homo rudolfensis* (such as OH 7 and 13, ER 1802); and early European fossils of archaic *Homo sapiens* (such as Mauer, Arago, Atapuerca). Selected comparative data on these fossils are included in Tables 1 and 2.

The Dmanisi mandible appears to be most similar to African *H. erectus* fossils, with fewer similarities to those from Zhoukoudian. It is clearly smaller than mandibles of earliest *Homo*, especially in the P₄ and M_{2,3} and in the corpus under the cheek teeth. The most striking feature of the tooth row is perhaps the great proportional decrease in tooth module ('area'), about 15–20% from M₁ to M₂ and from M₂ to M₃. No other specimen included in Table 2 presents such a strong distal reduction pattern, the closest approximations being OH 22 (from Bed IV, ~0.6 Myr, but lacking M₃ and all teeth mesial to P₃), Zhoukoudian G1-6 (~0.4 Myr) and Tighenif 3 (~0.7 Myr). If the total Zhoukoudian and Tighenif samples are considered, however, the first two molars are nearly equal in module, with M₃ smaller. Other *H. erectus* jaws, such as ER 992, Sangiran 1 and 9, and Zhoukoudian G1-7, have similar sized M₂ and M₃, usually with a smaller M₁, whereas the small M₁ of Mauer and Arago sets them apart from Dmanisi. The Thomas Quarry 1 mandible (~0.4 Myr) has very large mesial molars with a greatly reduced M₃. The Dmanisi M₃ (and to a lesser degree M₂) tilt lingually in their sockets, as seen also in ER 992 and WT 15000. The anterior teeth appear morphologically similar to those of African and Asian *H. erectus* (*sensu lato*).

In terms of tooth size, the Dmanisi M₂ is most comparable to those in the ER 992, WT 15000, OH 22, Zhoukoudian, Tighenif 3 and Mauer jaws. The Dmanisi premolars are small within the comparative sample, and especially narrow; they conform most closely to those of OH 13, the Zhoukoudian mean, Tighenif 3 and Mauer. The P₃ lingual cusp is strongly reduced, as seen also in the Zhoukoudian unworn specimen Uppsala M3549 described by Zdansky. The Dmanisi canine is small, with the buccolingual diameter only slightly smaller than the mesiodistal; its closest match is with those from Zhoukoudian and, surprisingly, OH 13. The incisors are close in size to those of the young juvenile ER 820 and the Zhoukoudian mean, lacking the buccolingual dominance of the North African and European specimens.

The Dmanisi corpus (Table 1) increases in relative thickness from the symphysis distally, although it must be emphasized that submolar depth is estimated. This pattern is also seen in ER 730 and 992, OH 22, Sangiran 9 (but not 1), possibly at Zhoukoudian, and in some but not all of the Tighenif and Arago

jaws. The remaining Tighenif, Arago, Mauer and Sangiran specimens have a rather consistent relative thickness along the tooth row. The Dmanisi jaw is small; its symphysis is comparable in size, and also in overall morphology, to that of ER 730, as well as to Sangiran 1, Tighenif 2, Thomas Quarry, Arago 2 and perhaps Zhoukoudian H1. By contrast, the symphysis neither slopes backward as much as in the early African fossils, nor is there a well developed trigone as in early European specimens. The corpus under the premolars is most similar in size to those of OH 13, OH 22 and Arago 2, as well as to the subadult WT 15000. Under the molars, size similarity is greatest to Zhoukoudian H1, Thomas Quarry and WT 15000, which probably would have presented a deeper (and thus far larger) corpus as an adult.

There are thus both morphological and metrical similarities to fossils spanning a wide range of time and space. The Dmanisi jaw is distinctive in the great dental reduction distally, as well as in relatively small tooth and corpus size combined with apparently great corpus robusticity under the cheek teeth; in addition, the alveolar arcade is narrow and the beginning of the ramus is placed far anteriorly. Among the most similar specimens overall are certain Turkana Basin fossils (ER 730 and 992, WT 15000),

OH 22 (lacking M_3) and Zhoukoudian G1-6 (with a much larger corpus). Selected members of the Tighenif/Thomas, Arago/Mauer and Sangiran samples show specific similarities as well. The most reasonable interpretation of this jaw is that it belonged to a population of *H. erectus*, possibly foreshadowing (European) early *H. sapiens*.

The date provided by the faunal and geophysical evidence indicates that the Dmanisi population was one of the oldest human groups outside Africa, certainly the most ancient in western Eurasia. Dates for Chinese and Indonesian *H. erectus* have long been in doubt, but recent studies have suggested ages of ~1.1 Myr for the Lantian Gongwangling cranium⁶ and ~1.7 Myr for the oldest Sangiran and Mojokerto fossils⁷. Although there are potential problems with these dates, the Indonesian evidence especially suggests hominid presence at least between 1.5–1.0 Myr. A similar age is estimated for the 'Ubeidiya archaeological site in Israel⁸. Ages between 2 and 1 Myr have also been suggested for a variety of localities in Europe, but the most recent review of this work⁹ indicates that there is little support for demonstrated human occupation before ~0.5 Myr. The Dmanisi human mandible, attributed to *H. erectus* and dated around 1.8 Myr or slightly younger, might be

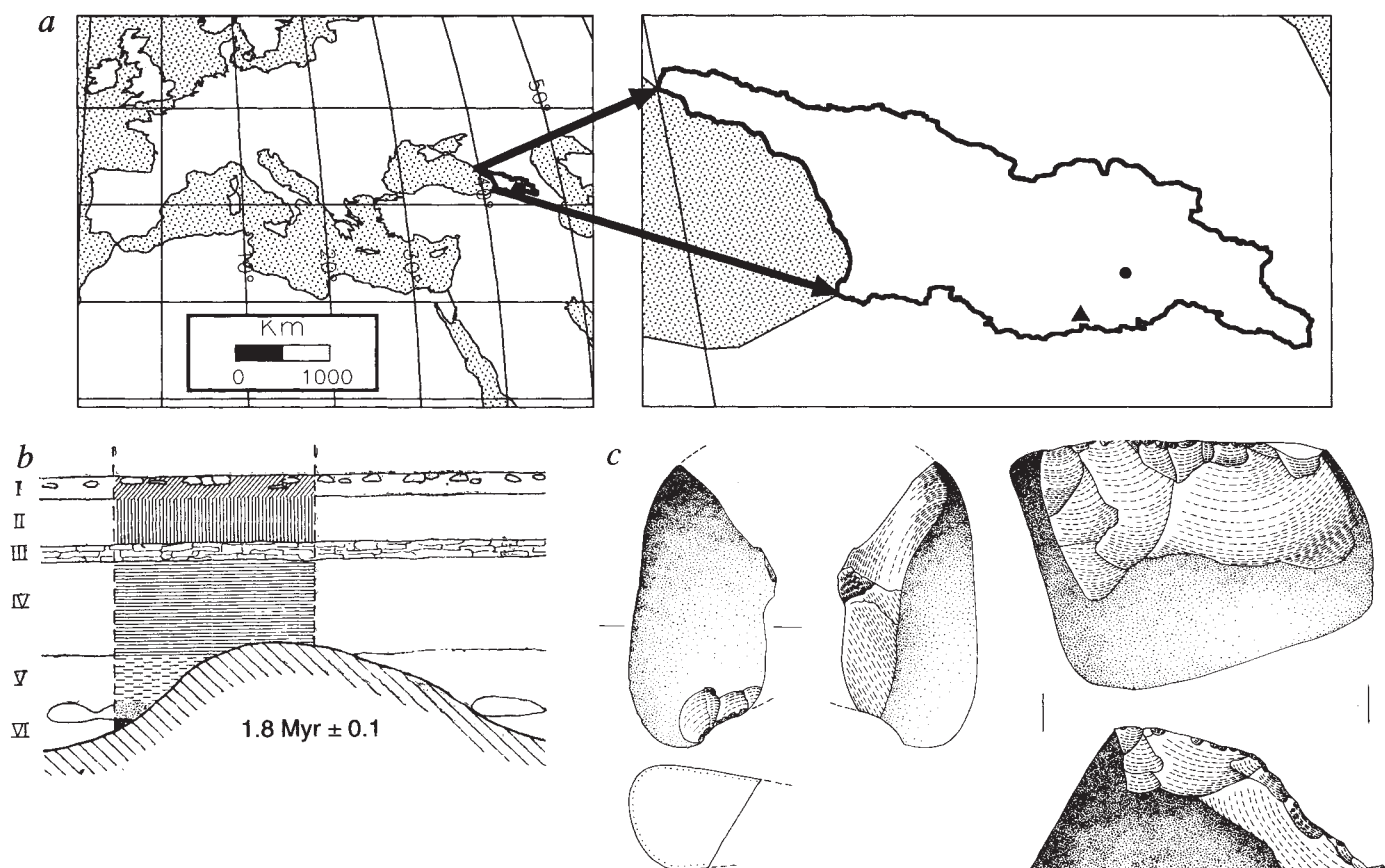


FIG. 1 a, Map showing rough location of Dmanisi (left, position of Georgia within Europe; right, Georgia, with Tbilisi (circle) and Dmanisi (triangle)); b, simplified stratigraphic column of hominid locality; and c, two artefacts found in the fossiliferous horizons (b and c after refs 2 and 15 respectively). In b, the sequence includes, from top downward: (I) grey clays with limestone blocks, of which the top formed the habitation surface for the mediaeval town (0.4 m); (II) yellow-brown loam with many artefacts but few bones (0.5 m); (III) thin limestone crust (0.2 m); (IV) brown loam with many fossil vertebrates and artefacts (1.5 m); (V) blackish-brown loam with thin sandy layers, many fossil vertebrates including the hominid mandible and numerous artefacts (1.0 m); (VI) compact blackish sand and volcanic ash with rare mammalian fossils; basalt with unaltered surface. In c, distance between vertical lines is 9.5 cm.

TABLE 1 Measurements (in mm) of the mandibular corpus in selected late Pliocene to late Middle Pleistocene *Homo* fossils

	Symphysis depth	Symphysis thickness	Symphysis ratio	P ₍₃₋₄₎ depth	P ₍₃₋₄₎ thickness	P ₍₃₋₄₎ ratio	M ₍₁₋₂₎ depth	M ₍₁₋₂₎ thickness	M ₍₁₋₂₎ ratio
Dmanisi 211 R	30.8	16.8	55%	26.8	18.5	69%	*24.7	18.4	74%
KNM ER 1802 R ¹⁰	36.0	24.5	68%	40.0	20.0	50%	38.0	27.0	71%
OH 13 R ¹⁰	25.0	18.0	72%	26.0	16.5	63%	28.5	22.5	79%
KNM ER 730 ¹⁰	32.5	17.5	54%	32.5	19.0	58%	31.5	19.0	60%
KNM ER 992 L ¹⁰	37.0	21.0	57%	31.0	20.0	65%	33.0	22.0	67%
KNM WT 15000 L ¹¹				27.2	18.1	67%	24.8	21.0	85%
OH 22 ¹⁰	33.5	20.0	60%	29.0	20.5	71%	28.5	21.0	74%
Sangiran 1b ¹⁰	32.0	17.0	53%	33.0	16.0	48%	33.0	17.0	52%
Sangiran 9 R ¹⁰	41.0	19.0	46%	39.0	21.5	55%	32.0	23.0	72%
Zhoukoudian G 1-6 ¹²	40.0	13.7	34%				34.0	17.3	51%
Zhoukoudian H 1 ¹²	32.3	14.0	43%				26.0	14.9	57%
Zhoukoudian Mean ¹²	34.5			28.5					
Tighenif 1 ¹³	37.5	18.5	49%	33.0	19.0	58%	35.0	22.0	63%
Tighenif 2 ¹³	31.0	17.0	55%	32.0	15.0	47%	31.5	17.5	56%
Tighenif 3 ¹³	35.0	19.0	54%	35.0	19.0	54%	38.0	21.0	55%
Thomas Quarry 1 ¹³				28.5	16.0	56%	26.5	*18.0	68%
Mauer	32.0	18.5	58%	32.5	18.5	57%	33.0	19.0	58%
Arago 2 ¹⁰	31.0	16.0	52%	30.5	17.0	56%	30.0	17.0	57%
Arago 13 ¹⁰	36.0	20.0	56%	31.0	22.5	73%	31.0	23.5	76%

Symphysis depth, minimum superoinferior distance between infradentale and base of symphysis (termed height in ref. 10, measurement no. 141); symphysis thickness, maximum mesiodistal (anteroposterior) distance perpendicular to preceding (termed depth, measurement no. 142); Symphysis ratio, percentage thickness/depth (rounded up); P₍₃₋₄₎ depth, minimum depth of corpus below septum between P₃ and P₄ or below middle of P₄, depending on authority (ref. 10, measurement no. 147, at P₄, lingual side; termed minimum height in ref. 13); P₍₃₋₄₎ thickness, maximum thickness of corpus perpendicular to preceding (ref. 10, measurement no. 148; minimum corpus breadth¹³). Ratio as for symphysis. M₍₁₋₂₎, as P₍₃₋₄₎ (ref. 10, measurement nos 154–155, at M₂; at M₂ in ref. 13). L or R following specimen identification refers to side of mandible when both are present; superscript number gives source of data.

* Value estimated on Dmanisi mandible (see Fig. 2g) or roughly estimated by authors from data in cited source.

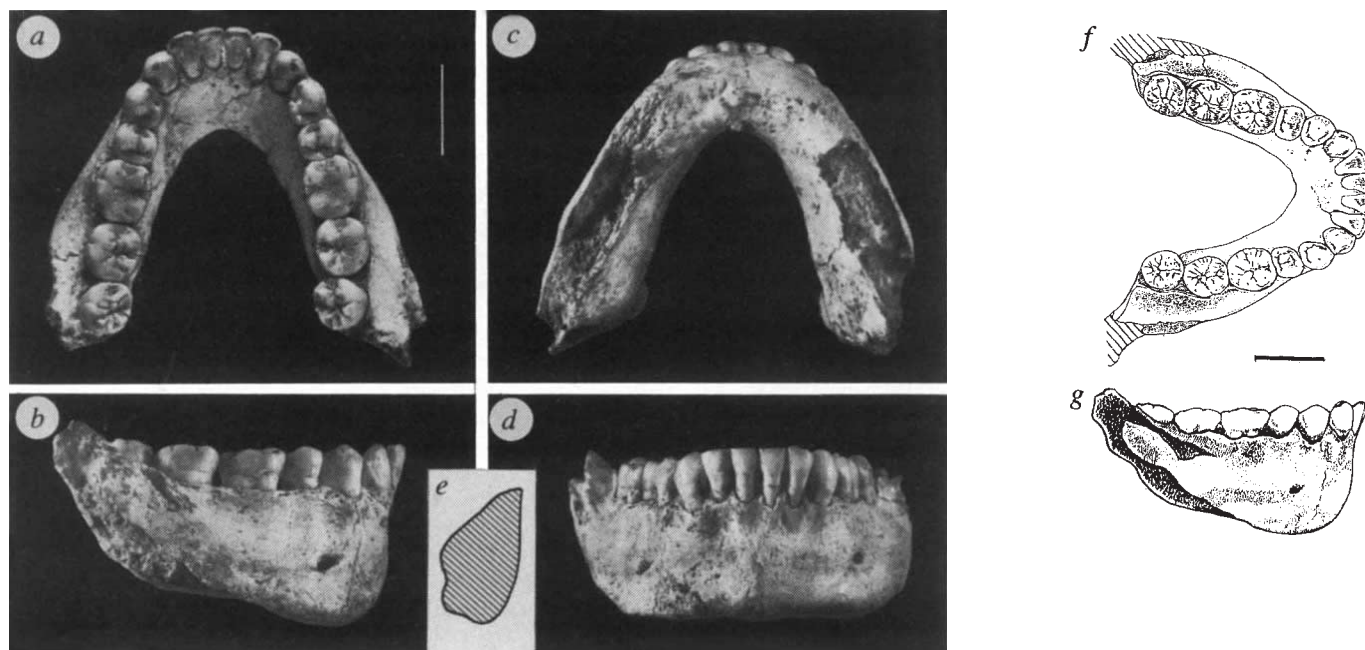


FIG. 2 Occlusal (a), right lateral (b), basal (c) and anterior (d) views of Dmanisi mandible, and symphyseal cross-section (e). Partially reconstructed drawing of Dmanisi mandible, in occlusal (f) and right lateral (g) views. Scale bars in a and f, 2 cm. The symphysis is nearly vertical in profile, with no sign of a mental protuberance, and it curves smoothly into the inferior border of the corpus. In anterior view, the shape appears subsquare. On each side a central swelling merges upward into the medioincisive and canine jugum, but does not extend anteriorly onto the symphysis; moreover, there is no incurvation below the incisors and thus no mental trigone. Internally, the laterally extensive alveolar planum slopes gently down to the superior transverse torus, but it does not extend posteriorly in a shelf. The genial fossa is distinct; the symphyseal base is flattened and thick anteroposteriorly, with weak digastric markings. The corpus is thick, and its superior and inferior borders may have been nearly parallel, although the base is broken below the molars (increasingly lacking posteriorly); anteriorly, the base is thick and slightly everted. A single large mental foramen is placed 14 mm below the mesial edge of P₄ on the right, whereas two smaller foramina occur on the left. A lateral prominence is developed below the anterior border of M₂, extending antero-inferiorly to merge with the well defined superior lateral torus. The alveolar arcade is broadly U-shaped, each set of cheek teeth curving inward at both

ends (the M₃s are most lateral). The retromolar space is small, whereas the extramolar sulcus is broad and hollowed. The anterior end of the buccinator groove lies at the level of mesial M₂, and the base of the triangular torus points superiorly, suggesting that the anterior border of the ramus might have been placed anteriorly and vertically. The slightly anteriorly inclined incisors show a small amount of flattened incisal edge wear. The canines are somewhat pointed, with a distinct external cingulum, but they do not project above the incisors. Both premolars bear distinct traces of cingulum and have a relatively large talonid. The P₃ is markedly asymmetric and rather triangular in outline, its buccal face being convex near the cervix. The anterior fovea is short and oriented nearly mesio-distally, whereas the posterior fovea is somewhat deeper and extends mesially to the very small lingual cusp. The crown of P₄ is almost quadrilateral in outline, wider than long and distinctly smaller (shorter) than that of P₃. The first and somewhat smaller second molars are quadrilateral in outline. The third molar is distinctly smaller, with a rather rounded outline. All three molars present a Y-6 cusp pattern, with the hypoconid surrounded by distinct grooves. The talonid of M₃ is somewhat longer than the trigonid and narrows distally; the occlusal surface of this tooth is tilted lingually by comparison to those of the other molars.

TABLE 2 Measurements (in mm) of the mandibular dentition in selected late Pliocene to late Middle Pleistocene *Homo* fossils

	I ₁ m-d	I ₁ b-l	I ₂ m-d	I ₂ b-l	C ₁ m-d	C ₁ b-l	P ₃ m-d	P ₃ b-l	P ₄ m-d	P ₄ b-l	M ₁ m-d	M ₁ b-l	M ₂ m-d	M ₂ b-l	M ₃ m-d	M ₃ b-l	Module ratios M ₁ /M ₂	M ₃ /M ₂	Molar Sizes
Dmanisi 211 R	5.9	5.8	6.6	6.4	8.7	8.2	9.0	9.2	8.1	9.2	13.2	12.3	12.3	11.5	11.2	10.6	115%	84%	1>2>3
Dmanisi 211 L	6.2	5.9	6.4	6.3	8.6	7.9	8.9	9.6	8.0	9.6	13.0	12.5	11.5	11.6	10.7	10.6	122%	85%	1>2>3
KNM ER 1802 R ¹⁰							10.7	11.5	11.4	12.0	14.8	13.0	16.6	14.2			82%		2>1
OH 7 L ¹⁰	6.5	6.7	7.3	7.6	8.9	10.1	9.5	9.7	10.4	10.7	14.1	12.5	15.7	13.7			82%		2>1
OH 13 R ¹⁰					7.6	7.9	9.0	8.7	9.0	9.9	13.0	11.6	14.2	12.0	14.8	12.4	88%	108%	3>2>1
KNM ER 820 L ¹⁰	6.1	6.3	6.3	6.9							12.5	10.6							
KNM ER 992 L ¹⁰			7.2	7.0	8.9	9.5	9.7	10.3	8.8	10.9	12.7	10.9	13.0	12.2	13.4	12.3	87%	104%	3>2>1
KNM WT 15000 L ¹⁴	6.6	6.8	7.5	8.3	8.9	9.6	8.9	10.1	9.0	10.1	12.4	11.0	12.5	12.0			91%		2>1
OH 22 ¹⁰							10.1	9.2	9.0	10.0	13.4	12.0	13.0	11.7			106%		1>2
Sangiran 1b ¹³									8.9	10.8	12.9	12.9	13.2	13.4	14.4	12.5	94%	102%	3>2>1
Sangiran 9 R ¹⁰					7.0	8.8	9.1	11.3	9.3	11.6			14.1	12.7	13.8	12.7		98%	2>3
Zhoukoudian G1-6 L ¹²	6.2	6.8		7.3	8.1	8.2	9.1	10.7	8.5	11.0	13.2	12.5	12.5	12.7	12.0	12.3	104%	93%	1>2>3
Zhoukoudian G1-7 ¹²													12.6	12.9	12.9	12.4		98%	2>3
Zhoukoudian Mean ¹²	6.4	6.4	6.8	7.0	8.7	9.4	8.7	9.9	9.0	9.8	12.5	11.8	12.5	12.1	11.7	11.3	98%	87%	2>1>3
Tighenif 1 ¹³			6.0	8.0			8.5	10.2	8.3	10.1	13.0	12.5	13.0	13.0	12.0	12.2	96%	87%	2>1>3
Tighenif 2 ¹³							8.6	11.1	8.8	11.0	13.9	12.8	14.1	13.3	13.4	12.5	95%	89%	2>1>3
Tighenif 3 ¹³			5.0	9.0	7.7	10.7	8.0	10.2	8.2	10.0	12.4	12.0	12.0	12.2	12.0	11.5	102%	94%	1>2>3
Thomas Quarry 1 ¹³									9.0	10.5	14.0	12.8	14.8	13.0	12.8	11.7	93%	78%	2>1>3
Mauer	5.5	7.5	6.9	7.7	7.4	8.9	8.1	9.0	7.5	9.2	11.3	11.2	12.7	12.0	10.9	11.9	83%	85%	2>3>1
Arago 13							10.1	11.0	9.5	12.9	13.6	13.6	14.6	14.0	13.4	13.0	90%	85%	2>1>3

Maximum mesiodistal (m-d) and buccolingual (b-l) dimensions of lower teeth. L or R following specimen identification refers to side of mandible when both are present; superscript number gives source of data. Module ratios represent products of m-d and b-l diameters of first and third molars, respectively, divided by those of second molars. ER 1802 is identified as *Homo rudolfensis*, OH 7 and OH 13 as *H. habilis*, Thomas Quarry, Mauer and Arago as early *H. sapiens*, other specimens as *H. erectus*, although ER 820-OH 22 are sometimes considered *H. ergaster*.

the oldest evidence of humans outside Africa. Its presence in Georgia further suggests that humans either waited outside the 'gates to Europe' for more than 1 Myr or inhabited the subcontinent at very low density during that interval. □

Received 11 August; accepted 1 December 1994.

1. Majsuradze, G., Pavlenshvil, E., Schmincke, H. & Sologashvili, D. *Jahrb. Römisch-German. Zentralmus.* **36**, 74–76 (1991).
2. Bosinski, G., Lordkipanidze, D., Majsuradze, G. & Tsvirelidze, M. *Jahrb. Römisch-German. Zentralmus.* **36**, 76–83 (1991).
3. Muskhelishvili, A. *Jahrb. Römisch-German. Zentralmus.* (in the press).
4. Gabunia, L. & Vekua, A. *Dmanisi Man and the Accompanying Vertebrate Fauna 1–60* (Metsniereba, Tbilisi, 1993).
5. Sologashvili, D., Majsuradze, G., Pavlenshvil, E. & Klopotovskaja, N. *Jahrb. Römisch-German. Zentralmus.* (in the press).
6. An, Z. & Ho, C. K. *Quat. Res.* **32**, 213–221 (1989).
7. Swisher, C. C. et al. *Science* **263**, 1118–1121 (1994).
8. Verosub, K. L. & Tchervov, E. in *Les Premiers Européens* (eds Bonifay, E. & Vandermeersch, B.) 237–242 (Comité Travaux Historiques Scientifiques, Paris, 1991).

9. Roebroeks, W. *Curr. Anthropol.* **35**, 301–305 (1994).
10. Wood, B. A. *Koobi Fora Research Project, vol. 4: Hominid Cranial Remains* (Clarendon, Oxford, 1991).
11. Walker, A. & Leakey, R. in *The Nariokotome Homo erectus Skeleton* (eds Walker, A. & Leakey, R.) 63–94 (Harvard Univ. Press, 1993).
12. Weidenreich, F. *Palaeontol. Sinica n.s.* **D1**, 1–180 (1937).
13. Rightmire, G. P. *The Evolution of Homo erectus* (Cambridge Univ. Press, 1990).
14. Brown, B. & Walker, A. in *The Nariokotome Homo erectus Skeleton* (eds Walker, A. & Leakey, R.) 161–192 (Harvard Univ. Press, 1993).
15. Bosinski, G., Bugianisvili, T., Mgeladze, N., Nioradze, M. & Tusabramisvili, D. *Jahrb. Römisch-German. Zentralmus.* **36**, 93–107 (1991).

ACKNOWLEDGEMENTS. We thank H. and M.-A. de Lumley for access to the rich comparative collections in the Laboratoire d'Anthropologie in Marseille, G. Bosinski (Römisch-Germanisches Zentralmuseum, Neuwied), E. Delson (Anthropologie, Lehman College of the City University of New York), J. L. Franzen (Forschungsinstitut Senckenberg, Frankfurt), F. C. Howell (Anthropology, University of California, Berkeley), G. P. Rightmire (Anthropology, State University of New York at Binghamton), A. Walker (Cell Biology and Anatomy, The Johns Hopkins University School of Medicine, Baltimore), T. White (Anthropology, University of California, Berkeley), D. Dean (Anatomy, Case Western Reserve University School of Medicine, Cleveland), H. Roth (Département d'Anthropologie et d'Ecologie, Université de Genève), B. Senut (Laboratoire d'Anthropologie du Muséum National d'Histoire Naturelle, Paris) for their help. E. Delson prepared the final versions of the manuscript and L. Meeker prepared Fig. 2, for which we are most grateful.

Mating cost of large floral displays in hermaphrodite plants

Lawrence D. Harder & Spencer C. H. Barrett*

Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, Canada

* Department of Botany, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S 3B2, Canada

HERMAPHRODITISM comprises outcrossing whenever the proximity of male and female organs allows self-fertilization¹ and interference between sexual functions². Many floral traits of animal-pollinated angiosperms encourage cross-fertilization³, as recognized by Darwin^{4–6}; however, these characteristics may also allow pollination between flowers on the same individual (geitonogamous self-pollination)^{7,8}. Simultaneous display of many flowers exemplifies this conflict. Although large floral displays promote outcrossing through enhanced pollinator attraction⁹, they could be costly in terms of lost mating opportunities^{10,11} if geitonogamy decreased outcrossed siring success by reducing pollen transfer between plants (pollen discounting¹²). We report here that, after manipulating the flower number of bee-pollinated *Eichhornia paniculata*

plants, we observed the predicted higher selfing and lower outcrossed siring success for larger inflorescences. Given the reduced fitness resulting when pollen receipt by one flower interferes with pollen export by another, we propose broadening traditional interpretations of floral design and display to recognize their roles in reducing geitonogamous pollen discounting.

To assess the effect of daily inflorescence size (3, 6, 9 or 12 flowers) on male- and female-fertility variation, we estimated the selfing frequency and outcrossed siring success in *Eichhornia paniculata* (Pontederiaceae), using allozyme markers. This species is annual, self-compatible, heterostylous and native to aquatic habitats in the Neotropics¹³. Each day, plants produce up to 20 flowers per inflorescence which open synchronously and last for about 6 h. Self and intramorph pollen have equivalent pollen-tube growth and siring ability and flowers seldom abort fertilized seeds⁷. As a result, fertility estimates are not biased by post-pollination discrimination against selfed progeny.

The experiment was conducted in Etobicoke, Ontario, and involved arrays of 35 or 36 long-styled plants (depending on treatments) with one inflorescence each. During a trial, an array included either plants of only one inflorescence size (pure arrays), or plants of two sizes (pairwise arrays). The experiment included all four pure arrays and all six possible pairwise arrays, with four replicates per combination. For pairwise arrays, we adjusted the number of plants for the two inflorescence-size treatments so that both were represented by the same total num-

ber of flowers (that is, equivalent pollen and ovule production per treatment). During the weeks before the experiment, we placed *E. paniculata* plants at the array locations so that resident bees (primarily *Bombus fervidus* and *B. vagans*) were familiar with this species.

For daily inflorescence size to affect reproductive success, the number of flowers visited per inflorescence by individual pollinators must increase with display size. Observations of bumblebee foraging during the experiment indicated that individual bees visited more flowers on larger inflorescences, although this response was context-dependent as the number of flowers visited on large inflorescences depended on the sizes of competing inflorescences in an array (Fig. 1a). The increase in number of flowers visited did not keep pace with increases in inflorescence size, so that the proportion of flowers visited per pollinator declined with inflorescence size. This decline was counter-balanced by the preference by pollinators for large inflorescences (Fig. 1b), so that the cumulative number of visits received per flower did not differ significantly among inflorescence size treatments ($F_{5,21} = 2.40$, $P > 0.05$).

Female fertility did not vary significantly with daily inflorescence size, as all inflorescence-size treatments produced equivalent numbers of seeds ($F_{3,17} = 0.11$, $P > 0.9$), regardless of the size of competing inflorescences (flower number $\times$ competitor size interaction, $F_{9,17} = 0.66$, $P > 0.7$). This uniform female success probably reflects the equal cumulative visitation of individual flowers within inflorescences of all sizes.

In contrast to female fertility, the incidence of self-fertilization varied positively with inflorescence size, so that selfing occurred twice as frequently in 12-flowered inflorescences as in 3-flowered inflorescences (Fig. 2a). The mechanism responsible for this increase can be inferred from the increase in selfing from bottom to top flowers within inflorescences (Fig. 2b). Such a pattern is consistent with geitonogamous movement of pollen by pollinators that habitually begin visiting inflorescences on bottom flowers and then move upwards^{7,8,14}, as did the bumblebees visiting *E. paniculata*⁷. In contrast, the alternative possibility that pollinators alter intrafloral self-pollination as they move up

inflorescences seems unlikely, because the duration of bee visits to flowers did not differ between the top and bottom halves of inflorescences (repeated-measures ANOVA, $F_{1,5} = 0.08$, $P > 0.75$; based on video-taped records of a related experiment). Hence, geitonogamous pollination is strongly implicated as the mechanism responsible for the pattern of selfing within inflorescences. This geitonogamy, coupled with the tendency for bees to visit more flowers on large inflorescences (Fig. 1a), also provides the most likely explanation for the observed positive relation of selfing to inflorescence size.

The increased selfing by large inflorescences was accompanied by reduced outcrossed siring success (Fig. 2c), indicating that pollen discounting varied directly with inflorescence size. Indeed, the smaller inflorescences in pairwise arrays realized a significant outcrossing advantage over larger inflorescences (paired *t*-test, $t_{23} = 2.74$, $P < 0.02$; mean $\pm$ s.e. difference in outcrossed siring success, 0.077 ± 0.028). We interpret this negative relation between selfing and outcrossed siring success as a direct outcome of geitonogamous pollination: such a mating cost associated with large floral displays has not previously been demonstrated empirically.

The occurrence of pollen discounting and its significance for the ecology and evolution of plants have been questioned^{11,15} because "[p]ollen grains are much more numerous than ovules, and . . . [c]onsequently, an increase in self-fertilization can often be achieved with minimal effect on the outcrossing pollen pool" (ref. 11). Although pollen grains are plentiful and individually inexpensive, the important consideration is the number of mating opportunities, rather than pollen production. Typically <1% of a plant's pollen reaches stigmas¹⁶, so we should ask how much of the pollen that might have reached other plants is used up in selfing, rather than merely how much pollen is involved in selfing. Unlike intrafloral selfing, geitonogamous self-pollination may often limit opportunities for outcrossing because it probably removes pollen directly from the restricted pollen pool that pollinators transport between flowers¹¹. As a result, geitonogamous pollen discounting should be a common consequence of displaying flowers simultaneously.

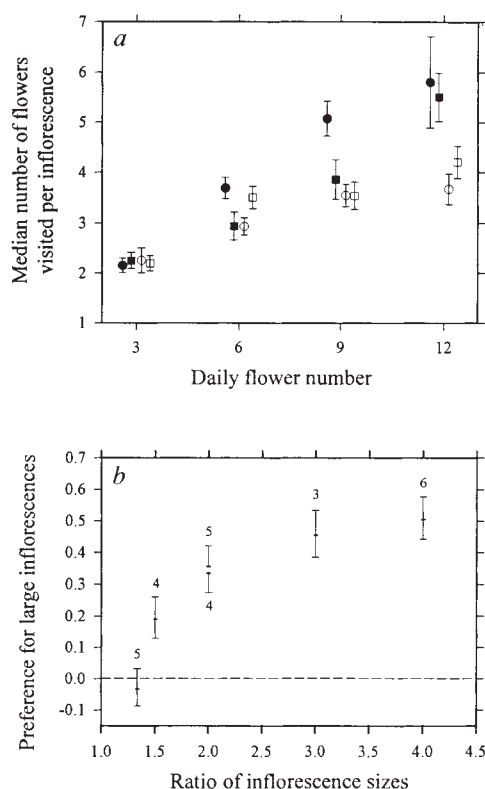


FIG. 1 The number of flowers per inflorescence affects the behaviour of pollinating bumblebees. In all trials we observed pollinator behaviour for 15 min during each of the first 3 h after flowering began. For each inflorescence visited by focal bees we recorded the inflorescence position in the array and number of flowers visited. a, Relation of the number of flowers visited per inflorescence to the size of visited and competing inflorescences. The median number of flowers visited per bee served as individual observations in a two-factor design with sizes of visited and competing inflorescences as classification variables. ANOVA of these data revealed a significant interaction in the effects of the sizes of the two inflorescence classes in an array ($F_{8,137} = 2.73$, $P < 0.01$): the general positive relation between the number of flowers visited and inflorescence size was most pronounced when competing inflorescences were small. Means and standard errors for 3 to 12 bees per treatment combination are shown. Symbols indicate size of competing inflorescences: solid circles, 3 flowers; solid squares, 6 flowers; open circles, 9 flowers; open squares, 12 flowers. b, Relative preference by bumblebees for large inflorescences in pairwise arrays. For focal bees that visited >10 inflorescences we estimated preference as $\ln\left(\frac{\text{odds that the bee visited a large inflorescence}}{\text{odds that any inflorescence was large}}\right)$, where the odds of an event is the ratio of the probability of the event occurring to the probability of the event not happening. The horizontal line indicates indifference, whereas positive preference values indicate preference for large inflorescences. Trend analysis²⁶ indicates that increases in the ratio of inflorescence sizes increased the preference by bees for large inflorescences (linear contrast, $F_{1,21} = 31.00$, $P < 0.001$), but reduced the incremental advantage of larger inflorescences (quadratic component of linear + quadratic contrast, $F_{1,21} = 9.04$, $P < 0.01$). In this analysis each observation was weighted by the inverse of the expected variance of the $\log_e$ of the numerator of the preference index²⁷ (the denominator is constant for trials involving the same combination of inflorescence sizes).

Our results provide insight into strategies for allocating flowering effort during the blooming period. Both inflorescence treatments in pairwise arrays had the same total number of flowers, which were deployed differently on individual plants. Such a situation occurs if plants produce the same numbers of flowers during their entire flowering period, but differ in their daily inflorescence sizes and thus flower for different periods. In our experiment, 'few-flowered' plants sired more outcrossed seeds than 'many-flowered' plants. Thus the strategy of presenting few flowers each day and flowering for a protracted period can maximize the outcrossing component of male fertility.

The mating implications of pollen discounting depend on the relative susceptibility of self pollen to post-pollination losses^{8,11}. If the plant is self-compatible and not subject to inbreeding depression, then discounting lowers the average heterozygosity of offspring by reducing the siring of outcrossed seeds. In contrast, if selfed seeds suffer inbreeding depression, geitonogamy reduces both the parent plant's female and male fertility. Self-incompatibility protects plants from losses of female fertility through selfing; however, it exacerbates male losses associated with pollen discounting because self pollen sires no seeds.

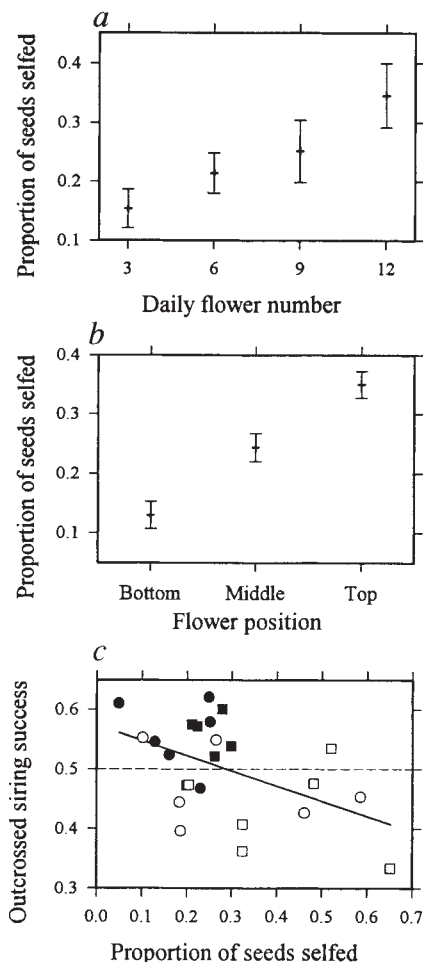
Pollen discounting bears two main consequences for the evolution of plant reproduction. First, discounting may influence the evolution of self-fertilization. With no discounting, selfing indi-

viduals realize a fitness advantage, as they contribute an average of three gene copies to the next generation (two through each selfed seed and one through paternal contributions to outcrossed seed) for every two copies contributed by outcrossing individuals¹⁷. Discounting diminishes this advantage because gametes involved in selfing reduce the number of gametes available for outcrossing^{12,18}. Recognition of pollen discounting as a potential constraint on the evolution of selfing has prompted recent attempts to measure discounting, with differing results^{15,19,22}. In contrast to earlier work, our study provides the first demonstration of a potentially widespread mechanism for discounting, namely geitonogamy. This mode of discounting may explain the disparity between previous studies, as studies of multiflowered plants demonstrated discounting^{19,20,22}, whereas those of single-flowered plants did not^{15,21}.

Lost outcrossing opportunities associated with large floral displays may also promote the evolution of floral mechanisms that decouple the benefits of large displays for pollinator attraction from the costs of geitonogamous pollen discounting. Consider two examples. First, when male and female functions are temporally separated in individual flowers (dichogamy) and inflorescence structure causes pollinators to visit female- before male-phase flowers, inflorescences can display many flowers without causing much geitonogamy²³. This interpretation may explain

FIG. 2 Daily inflorescence size influences the trade-off between selfing and outcrossed siring success. Plants in all arrays contained known electrophoretic markers, which allowed us to identify the inflorescence size of the father of each seed and to quantify the fraction of seeds that was self-fertilized for each inflorescence-size treatment (s_i and s_j , respectively for the two treatments in an array). **a**, Relation of the proportion of selfed seeds (mean $\pm$ s.e.) to inflorescence size. The frequency of selfing (s_i) increased linearly with increasing inflorescence size (test for linear trend, $F_{1,16} = 9.47$, $P < 0.01$); this result did not vary with the size of competing inflorescences in the array (flower number $\times$ competitor size interaction, $F_{9,16} = 0.51$, $P > 0.8$). **b**, Relation of the proportion of selfed seeds (mean $\pm$ s.e.) to the position of a flower within an inflorescence. Self-fertilization (s_i) increases significantly from bottom to top flowers ($F_{2,32} = 23.75$, $P < 0.001$), regardless of inflorescence size (flower position $\times$ inflorescence size interaction, $F_{6,32} = 0.99$, $P > 0.4$), presumably because of geitonogamous pollination resulting from the vertical structure of *E. paniculata* inflorescences and the tendency of bees to move upwards on inflorescences⁷. **c**, Dependence of relative outcrossed siring success on the frequency of selfing (s_i) in pairwise arrays. Outcrossed siring success (O_i) is the fraction of all outcrossed seeds per array that were sired by inflorescence treatment i , $O_i = (1 - s_i)T_i - B_i + B_j/(1 - s_i)T_i + (1 - s_j)T_j$, where: T_i and T_j are the total numbers of seeds produced by the two inflorescence-size treatments in an array, B_i is the number of seeds produced by treatment i that were sired by treatment j , and B_j is the number of seeds sired by treatment j on treatment i plants. In the absence of pollen discounting, both treatments in an array should realize equivalent outcrossed siring success (that is, $O_i = 0.5$; dashed line). Instead, the negative relation between outcrossed siring success and selfing ($\hat{O}_i = 0.573 - 0.254s_i$; solid line) indicates significant pollen discounting ($F_{1,22} = 18.43$, $P < 0.001$, $R^2 = 0.239$). Pure arrays were excluded from this analysis because outcrossing is constrained to be the exact complement of selfing when only one treatment is involved. Furthermore, we included data for only one inflorescence treatment per array (randomly selected) to avoid the lack of independence caused by the siring success of one treatment being the complement of that for the competing treatment in an array. Symbols indicate inflorescence sizes: solid circles, 3 flowers; solid squares, 6 flowers; open circles, 9 flowers; open squares, 12 flowers.

METHODS. In pairwise arrays, plants with one inflorescence size were homozygous for the fast AAT-3 allele whereas plants in the competing treatment were homozygous for the slow AAT-3 allele. Pure arrays contained equal numbers of plants with one of these homozygous AAT-3 genotypes. Plants in all arrays were polymorphic for *PGI-2*. Once flowers wilted after an experimental trial, one flower from each of the top, middle and bottom thirds of the inflorescence was marked with paint. Capsules produced by marked flowers were collected after the seeds had ripened (11–12 days) and stored in separate envelopes until the seeds could be counted and assayed electrophoretically. On the basis of the genotypes of three randomly selected seeds per capsule (that is,



nine seeds per plant) we estimated the fraction of selfed seed (s_i) and the associated standard error (based on 100 bootstrap samples) according to flower position and inflorescence size with Ritland's²⁸ multilocus selfing rate program. For **a** and **b**, the influences on s_i were assessed with a repeated-measures ANOVA in which individual estimates of s_i were weighted by the inverse of their squared standard error²⁷. Supplemental hand pollination indicated that seed production was not resource limited⁷.

the paradox that although dichogamy is traditionally interpreted as an anti-selfing mechanism², many dichogamous species also possess physiological self-incompatibility²⁴. The apparent redundancy of two mechanisms that prevent selfing² is resolved by recognizing inter-floral dichogamy as a mechanism that promotes outcrossed siring success by limiting pollen discounting, a role that self-incompatibility cannot serve. The second example involves heterostyly, a genetic polymorphism in which plants of each morph have anthers and stigmas in dissimilar positions, with the other morph(s) producing the reciprocal arrangement(s). The pronounced separation of sex organs in heterostylous plants reduces pollen transfer within and between plants of the same morph²⁵ and, consequently, more pollen remains on the pollinator until it visits the alternative morph(s). Like dichogamy, heterostyly may function, in part, to reduce the mating cost of large floral displays. However, whereas dichogamy typically functions within inflorescences, heterostyly limits pollen dispersal between inflorescences of the same morph. These examples imply that the negative relation between selfing and outcrossed siring success demonstrated by our experiment may be an important, though largely unexplored, influence on the evolution of floral design and display. □

Received 22 June; accepted 14 December 1994.

1. Jarne, P. & Charlesworth, D. A. *Rev. Ecol. Syst.* **24**, 441–446 (1993).
2. Lloyd, D. G. & Webb, C. J. *New Zeal. J. Bot.* **24**, 135–162 (1986).
3. Richards, A. J. *Plant Breeding Systems* (Allen and Unwin, London, 1986).

4. Darwin, C. R. *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom* (Murray, London, 1876).
5. Darwin, C. R. *The Different Forms of Flowers on Plants of the Same Species* (Murray, London, 1877).
6. Darwin, C. R. *The Various Contrivances by Which Orchids are Fertilised by Insects* 2nd edn (Murray, London, 1877).
7. Barrett, S. C. H., Harder, L. D. & Cole, W. W. *Funct. Ecol.* **8**, 526–535 (1994).
8. de Jong, T. J., Waser, N. M. & Klinkhamer, P. G. L. *Trends Ecol. Evol.* **8**, 321–325 (1993).
9. Klinkhamer, P. G. L., de Jong, T. J. & de Bruyn, G.-J. *Oikos* **54**, 201–204 (1989).
10. Charlesworth, D. & Charlesworth, B. *Evolution* **41**, 948–968 (1987).
11. Lloyd, D. G. *Int. J. Pl. Sci.* **153**, 370–380 (1992).
12. Holsinger, K. E., Feldman, M. W. & Christiansen, F. B. *Am. Nat.* **124**, 446–453 (1984).
13. Barrett, S. C. H. *Biol. J. Linn. Soc.* **25**, 41–60 (1985).
14. Crawford, T. J. in *Evolutionary Ecology* (ed Shorrocks, B.) 135–173 (Blackwell Scientific, London, 1984).
15. Holsinger, K. E. & Thomson, J. D. *Am. Nat.* **144**, 799–812 (1994).
16. Harder, L. D. & Barrett, S. C. H. in *Floral Biology: Studies on Floral Evolution in Animal-Pollinated Plants* (eds Lloyd, D. G. & Barrett, S. C. H.) (Chapman and Hall, New York, in the press).
17. Fisher, R. A. *Ann. Eugenics* **11**, 53–63 (1941).
18. Nagylaki, T. J. *theor. Biol.* **59**, 55–58 (1976).
19. Holsinger, K. E. in *Ecology and Evolution of Plant Reproduction* (ed. Wyatt, R.) 169–191 (Chapman and Hall, New York, 1992).
20. Kohn, J. R. & Barrett, S. C. H. *Evolution* (in the press).
21. Rausher, M. D., Augustine, D. & VanderKooi, A. *Evolution* **47**, 1688–1695 (1993).
22. Ritland, K. *Am. Nat.* **138**, 1049–1057 (1991).
23. Klinkhamer, P. G. L. & de Jong, T. J. *Oikos* **66**, 180–184 (1993).
24. Bertin, R. I. *Am. J. Bot.* **80**, 557–560 (1993).
25. Lloyd, D. G. & Webb, C. J. in *Evolution and Function of Heterostyly* (ed. Barrett, S. C. H.) 179–207 (Springer, Berlin, 1992).
26. Kirk, R. E. *Experimental Design* 2nd edn (Brooks/Cole, Belmont, 1982).
27. Neter, J., Wasserman, W. & Kutner, M. H. *Applied Linear Statistical Models* 3rd edn (Irwin, Homewood, 1990).
28. Ritland, K. *J. Hered.* **81**, 235–237 (1990).

ACKNOWLEDGEMENTS. We thank W. W. Cole and F. Strumas for technical assistance and D. Charlesworth, C. G. Eckert, S. W. Graham and B. E. Lyon for comments on the manuscript. This research was funded by Research Grants from the Natural Sciences and Engineering Research Council of Canada.

Dynamics of neuronal interactions in monkey cortex in relation to behavioural events

E. Vaadia, I. Haalman, M. Abeles, H. Bergman, Y. Prut, H. Slovin & A. Aertsen*

Department of Physiology, Hadassah School of Medicine, and the Center for Neural Computation, The Hebrew University, Jerusalem 91010, Israel

* The Center for Research of Higher Brain Functions, The Weizmann Institute of Science, Rehovot 76100, Israel

It is possible that brain cortical function is mediated by dynamic modulation of coherent firing in groups of neurons. Indeed, a correlation of firing between cortical neurons, seen following sensory stimuli or during motor behaviour, has been described^{1–5}. However, the time course of modifications of correlation in relation to behaviour was not evaluated systematically. Here we show that correlated firing between single neurons, recorded simultaneously in the frontal cortex of monkeys performing a behavioural task, evolves within a fraction of a second, and in systematic relation to behavioural events. Moreover, the dynamic patterns of correlation depend on the distance between neurons, and can emerge even without modulation of the firing rates. These findings support the notion that neurons can associate rapidly into a functional group in order to perform a computational task, at the same time becoming dissociated from concurrently activated competing groups. Thus, they call for a revision of prevailing models of neural coding that rely solely on single neuron firing rates^{6–8}.

The neuronal activity of 6–16 neurons in the frontal cortex of a rhesus monkey was recorded simultaneously while the animal performed a spatial delayed response task (Fig. 1). Interactions between pairs of neurons were studied by dynamic cross-correlation

(joint peri-stimulus time histogram, JPSTH⁹). This analysis highlights the detailed time structure of the correlation of firing between two neurons and its time-locking to a third event, such as the onset of a stimulus or the initiation of a movement. All JPSTHs were normalized to remove the contributions of stimulus- (or movement-) induced co-variation of single-neuron firing rates. Thus, the normalized JPSTH reflects the net effect of direct and indirect interactions of the two neurons, expressed as correlation coefficients⁹. We also computed the relative deviation of the observed coincidence rate from the expected rate for independently firing neurons. Its maximal absolute value, measured along the coincidence-time histogram (the main diagonal of the JPSTH; Fig. 2) and expressed as a percentage of the expected rate, was denoted as the 'maximal modulation depth'.

Figure 2 shows JPSTHs for two neurons in the premotor cortex for two behavioural paradigms. The peri-stimulus time (PST) histograms for paradigms GO (Fig. 2a) and NO-GO (Fig. 2b) are very similar, and show very little (x-axis) or no (y-axis) modulation of the firing rates. Thus, PST analysis indicates that these two neurons, taken individually, neither respond to the ready signal nor contribute to the discrimination between the two behavioural paradigms. Further, the cross-correlograms in Fig. 2a and b, which both have a wide central peak, are very similar. Thus, firing rates and time-averaged correlation do not discriminate between the two behavioural paradigms.

Such discrimination can be made, however, on examining the dynamics of the correlation. The main diagonal (bottom left to upper right) in each colour-coded matrix, emphasized in the coincidence-time histogram (Fig. 2), reveals strong modulations of co-firing between these two neurons, which are temporally linked to the ready signal. Moreover, the evolution of co-firing is virtually opposite in the two behavioural conditions: during performance of the GO paradigm (Fig. 2a) high co-firing occurs only during the first half of the time-window, whereas in NO-GO trials (Fig. 2b) it occurs in the second half. Note that in Fig. 2a (GO) the co-firing is maximal just after the ready signal, whereas in Fig. 2b (NO-GO) the co-firing at that time falls to zero. We stress that this difference in the evolution of co-firing

could not be predicted from the responses of the two neurons alone, nor was it reflected in the time-averaged cross-correlograms. However, the dynamics of the correlation reveal that these two neurons participate in the coding of the sensory event and its behavioural context.

Other examples of correlation dynamics are shown in Fig. 3. Neurons 0 and 1 were recorded by a single microelectrode. We refer to them as neighbouring neurons. The second pair consists of neuron 0 (from the first pair) and neuron 6, recorded by the tip of a different electrode, located 400 μm deeper and 500 μm lateral to the first electrode. Hence, we refer to them as distant neurons.

Figure 3a, b shows JPSTHs for the two neighbouring neurons (0, 1) during an interval of 3 seconds surrounding the initiation of saccadic eye movements to the right (Fig. 3a) and to the left (Fig. 3b). The PST histograms show that the firing rates of the neurons were modulated just before and during the saccades. The high peaks in the cross-correlograms indicate that, on average, the amount of correlation between the two neurons is similar and strong in both conditions. The JPSTH analysis, however, clearly reveals that these time-averaged cross-correlations are misleading: whereas the average correlation is indeed similar, dramatic modifications of co-firing occur near the initiation of the saccades. Co-firing increases just as the rightward saccade starts, and reaches its highest level within a few tens of milliseconds. During this brief period (white cluster in matrix a) the neurons co-fired on average nine times per saccade, about twice the expected number (maximal modulation depth equals 110%). The opposite change takes place for leftward saccades (Fig. 3b), where the correlation decreases to zero (no-correlation) when the saccade starts.

The JPSTHs in Fig. 3c and d show the same analysis for the distant pair (0, 6). Here, the average correlation is negative. Such

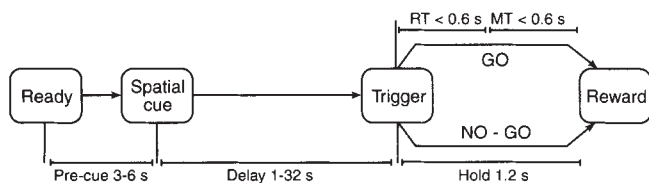


FIG. 1 The spatial delayed-response task. Trials were initiated when the monkey touched a central key and a central red light was turned on ('ready'). After a variable delay (3–6 s), one of two target keys was illuminated for 200 ms ('cue'). Then, after a delay of 1–32 s, the central red light was dimmed ('trigger'), instructing the monkey to exhibit the appropriate behavioural response. In one paradigm ('GO'), the monkey was rewarded for releasing the centre key within 0.6 s (RT, reaction time) after the trigger signal, and touching the correct target key within the next 0.6 s (MT, movement time). In the second paradigm ('NO-GO'), the monkey was rewarded for continuing to touch the centre key for at least 1.2 s after the trigger. Special peripheral lights instructed the monkey to switch paradigms after every fourth correct trial.

METHODS. Two monkeys (*Macaca mulatta*, female, 3–4 kg) were trained to perform the task. Recording sessions began after the monkey was well trained. In each session we simultaneously recorded the activity of 6–16 neurons during the performance of 400–800 correct trials, with 100–200 alternations between the two behavioural paradigms. Six microelectrodes (glass-coated tungsten, 0.2–1.5 M Ω impedance) were inserted into the frontal cortex in a circular array, with 500–1,000 μm inter-electrode distance. Spike-sorters were used to isolate the activity of 1–3 single units from each electrode. Typically, in each session we recorded 6–10 well-isolated units and 6–10 clusters with a mixture of 2–3 spike-shapes in each cluster. Eye movements were monitored by Ag–AgCl cup-electrodes. The locations of the electrodes' tracks in the cortex were reconstructed by standard histological techniques. Statistical significance of JPSTH correlation coefficients was evaluated using the 'surprise'-measure²³. Statistical tests for the time-averaged cross-correlograms were described by Abeles²⁴. All procedures were carried out in accordance with the NIH and Hebrew University regulations for animal care.

correlation can result from reciprocal effects of correlated inputs to the two cells (while these inputs excite one neuron, they inhibit the other). Alternatively, it may reflect mutual inhibition between the two. Again, the JPSTH shows that the average correlation is misleading. In fact, the correlation is only clearly negative just after the initiation of rightward saccades (with a

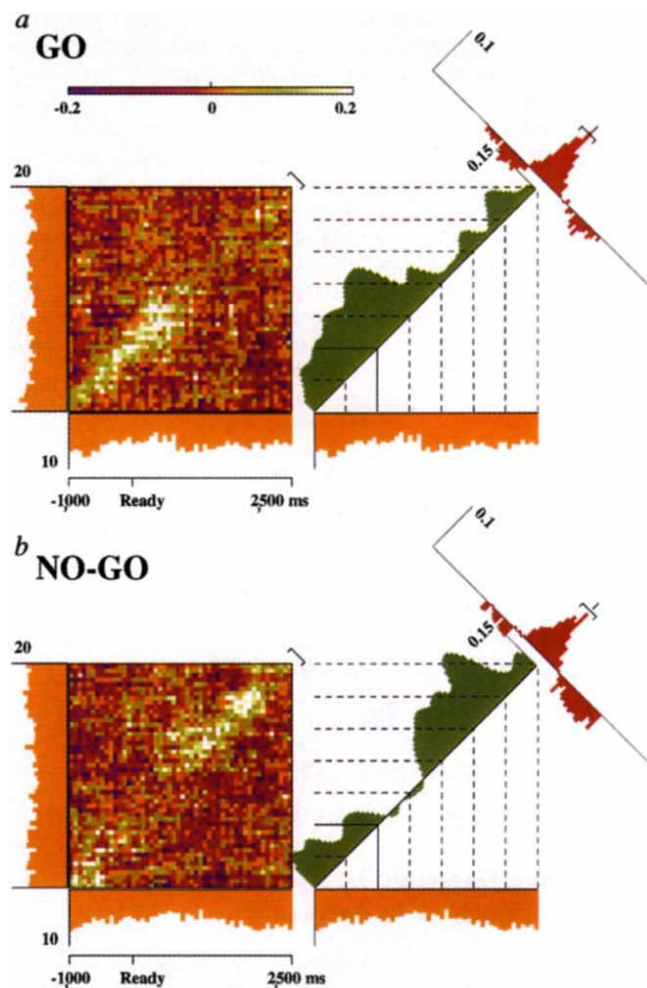


FIG. 2 Dynamic modification of correlated firing of two frontal cortex neurons in relation to the 'ready' signal. a, JPSTH constructed from 221 GO trials, with 4,414 spikes of neuron 1 and 10,121 spikes of neuron 2. Observe the gradual build-up of excess correlation, which starts before the ready signal, reaches a peak 400 ms after the signal (at that point the maximal modulation depth is 50%) and then decays. b, JPSTH constructed from 194 NO-GO trials for the same neurons (neuron 1: 3,764 spikes; neuron 2: 8,715 spikes). Note the different temporal evolution of the correlation compared with a. Each JPSTH is composed of four displays. (1) A colour-coded matrix of correlation coefficients (left half of plots a and b), with each bin depicting the amount of correlation of the two neurons at the corresponding pair of time delays, relative to the ready signal; colour scale (shown in a) from blue (minimum) to white (maximum). (2) Peri-stimulus time histograms (PSTHs; in orange), along the x-y axes of the matrix, depicting the firing rates of the two neurons relative to the time of the ready signal; x-axis: neuron 1, full scale, 10 spikes per s, y-axis: neuron 2, full scale: 20 spikes per s (the PSTH of neuron 1 is duplicated along the x-axis in the right-hand plot). (3) Coincidence-time histogram (in green; full scale, 0.15) showing the temporal evolution of the correlation coefficient relative to the ready signal. This histogram measures excess correlation within the time-window marked by the] sign at the upper-right corner of the matrix. (4) The conventional, time-averaged cross-correlogram (in red; full scale, 0.1). Grid line intervals, 500 ms. The timing of the ready signal is marked by solid grid lines. Bin width, 70 ms (PSTHs) and 70 ms $\times$ 70 ms (matrix).

maximal modulation depth of 70%), whereas it is even slightly positive near leftward saccades (Fig. 3*d*). Similar dynamics of negative correlation were found between neurons 1 and 6 (not shown).

Again, these changes of correlation could not be predicted from the neurons' firing rates. The correlation either increased (Fig. 3*a*) or decreased (Fig. 3*c*) near the saccade, whereas the firing rates of the neurons increased, regardless of saccade direction.

Analysis of 947 neuronal pairs revealed correlated activity in 499 pairs. Analysis of all 499 pairs in detail using the JPSTH method revealed that modulation of correlation between neurons is quite frequent; rapid changes of the rate of co-firing in relation to behaviour were observed in 308/499 cases (61%), with time constants of modulation down to a few tens of milliseconds.

The time-averaged cross-correlograms of these pairs had a peak centred around time delay 0 ('positive' correlation), a trough centred around 0 ('negative' correlation), or an asymmetrical pattern around 0, which included both a peak and a trough ('compound' correlation). The peaks and troughs decayed within tens to a few hundred milliseconds (Figs 2 and 3). To test whether the pattern of correlation depends on the distance between neurons, as shown in Fig. 3, we compared the distribution of correlation patterns for neuronal

pairs recorded by the same electrode and the correlations of pairs recorded by different electrodes (Fig. 4). A positive time-averaged correlation was observed between neighbouring neurons as well as between distant neurons, but more frequently between neighbouring ones. By contrast, a negative correlation was found only between distant neurons, and never between neighbouring ones.

Our results show that many cortical neurons exhibit rapid modulations of discharge correlation in relation to behavioural events. Epochs with a particular correlation may last from few tens of milliseconds to several seconds. The observed modulations of correlation may be, but do not have to be, associated with changes in the individual neurons firing rates. These findings support the notion that a single neuron can intermittently participate in different computations by rapidly changing its coupling to other neurons, without associated changes in firing rate.

Previous correlation studies concentrated mainly on relatively precise coincidences, with a resolution of only few milliseconds¹⁰⁻¹², the assumed time jitters of direct synaptic interactions. Here we chose to describe the phenomenon of loose synchrony, also measured in the visual cortex¹³. As we described elsewhere¹⁴, however, precise coincidences do occur in these data. For example, JPST analysis of the recordings in Fig. 2*a*, using a narrower bin of 5 ms, revealed that the modulation depth

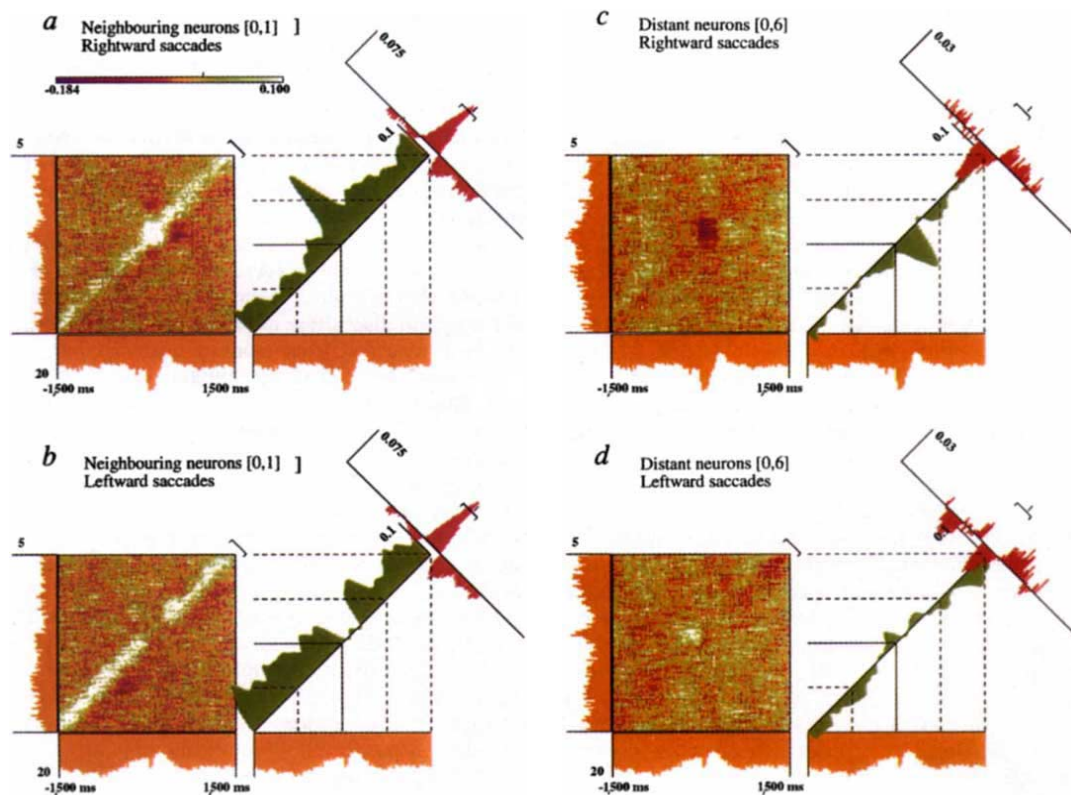


FIG. 3 Dynamic modification of correlated firing between neurons in the frontal eye field in relation to the onset of saccadic eye movements. *a, b*, JPSTHs for neurons (0, 1), recorded by one microelectrode ('neighbouring' neurons). *c, d*, JPSTHs for neurons (0, 6). Neuron 0 is from the first pair, neuron 6 was recorded by another electrode ('distant' neurons). The following features are apparent. First, the averaged cross-correlograms (in red) for each pair are very similar. However, the correlation dynamics are temporally linked to the saccades and depend strongly on their direction, as shown by the matrix and the (green) coincidence-time histograms (*a* compared with *b* and *c* compared with *d*). Second, the time-averaged correlation between neighbouring neurons (*a, b*) is positive (that is, the probability that either of the neurons will fire a spike is higher around the times the other neuron fires),

whereas the correlation between distant neurons (*c, d*) is negative. Third, the temporal changes of correlation could not be predicted from the firing rates of the two neurons. The correlation either increased (*a*) or decreased (*c*) near the time of saccade initiation, whereas the firing rates of both neurons increased around the onset of saccades, regardless of its direction. The normalization and format of the JPSTHs are the same as in Fig. 2. Bin size, 30 ms. The JPSTHs around onsets of rightward saccades (*a, c*) were constructed from 776 saccade, 33,882 spikes of neuron 0 (*a, c*), 4,299 spikes of neuron 1 (*a*) and 6,927 spikes of neuron 6 (*c*). The JPSTHs in *b* and *d* were constructed from 734 saccades, 32,621 spikes of neuron 0 (*b, d*), 4,167 spikes of neuron 1 (*b*) and 5,992 spike of neuron 6 (*d*).

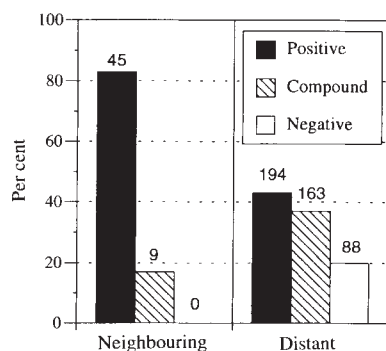


FIG. 4 Distribution of cross-correlogram patterns in neighbouring and distant neurons. Time-averaged cross-correlograms between pairs of neighbouring neurons are, in most cases, characterized by positive peaks. Note that negative cross-correlograms are completely missing in the population of neighbouring neurons. The number of neuronal pairs in each class is given above the bars. The difference between the two populations is statistically significant at a level of 0.001 ($\chi^2 = 318.77$).

of precise coincidences (with jitters of only ± 2.5 ms) reached values up to 620%.

The wide peaks and troughs (tens to hundreds of milliseconds wide) indicate that the time-constraints of the underlying processes are loose. The mechanisms of such dynamic correlation (including precise coincidences) are unknown. The correlation could emerge from repeated volleys of direct synaptic interactions between the neurons; or, more probably, it could arise from changes in the pattern of activity of a large number of neurons, interacting with the sampled neurons in a correlated manner. Regardless of the mechanism, however, the modifications of correlation between two neurons in relation to stimulation and behaviour most probably reflect changes in the organization of spike activity in larger groups of neurons. Indeed, recent model studies have shown that similar dynamic organization can be accomplished in large networks, even without associated modifications of the synaptic weights¹⁵⁻¹⁷.

Our results also support and extend anatomical and physiological findings indicating that functional groups are organized in spatial clusters^{3,11,18-22}. Our findings suggest that neighbouring neurons tend to share common inputs of the same sign (either inhibitory or excitatory), whereas the effects on more distant neurons (in our study 500–1,000 μ m apart) are mixed. Therefore, when the common drive is increased, neighbouring neurons tend to be activated in unison, enabling them to operate as a coherent functional group for a short while. Concurrently, the negative correlation between neurons in one group and those in another, more distant group can accentuate the demarcation among groups. Thus, the spatio-temporal organization of activity in the network allows for the rapid association of neurons into a functional group, at the same time dissociating such a group from concurrently activated, competing groups. $\square$

11. Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. *J. Neurosci.* **6**, 1160–1170 (1986).
12. Fetz, E. E., Toyama, K. & Smith, W. in *Cerebral Cortex* Vol. 9 (ed. Peters, A.) 1–47 (Plenum, New York, 1991).
13. Nelson, J. I., Salin, P. A., Munk, M. H. J., Arzi, M. & Bullier, J. *Vis. Neurosci.* **9**, 21–37 (1992).
14. Abeles, M., Prut, Y., Bergman, H., Vaadia, E. & Aertsen, A. in *Brain Theory: Spatio-Temporal Aspects of Brain Function* (ed. Aertsen, A.) 149–181 (Elsevier, Amsterdam, 1993).
15. Aertsen, A., Erb, M. & Palm, G. *Physica D* **75**, 103–128 (1994).
16. Hansel, D. & Sompolinsky, H. *Phys. Rev. Lett.* **68**, 718–721 (1992).
17. Kaneko, K. *Physica D* **41**, 137–172 (1990).
18. Gray, C. M. & Singer, W. *Nature* **338**, 334–337 (1989).
19. Kruger, J. *Vis. Neurosci.* **5**, 135–142 (1990).
20. Selemon, L. D. & Goldman-Rakic, P. S. *J. comp. Neurol.* **297**, 359–376 (1990).
21. Michalski, A., Gerstein, G. L., Czarkowska, J. & Tarnecki, R., *Expl Brain Res.* **51**, 97–107 (1983).
22. Thomson, A. M. & Deuchars, J. *Trends Neurosci.* **17**, 119–129 (1994).
23. Palm, G., Aertsen, A. M. H. J. & Gerstein, G. L. *Biol. Cybernet.* **59**, 1–11 (1988).
24. Abeles, M. *J. Neurosci. Meth.* **5**, 317–325 (1982).

ACKNOWLEDGEMENTS. We thank H. B. Barlow for comments on an earlier version of the manuscript, and G. Mato for discussion. The study was supported in part by grants from the Israeli Academy of Sciences and Humanities, the American-Israel Bi-national Science Foundation, the German 'Bundesministerium für Forschung und Technologie (BMFT)' and the Human Frontier Science Program (HFSP).

Neuronal correlates of inferred motion in primate posterior parietal cortex

John A. Assad & John H. R. Maunsell

Division of Neuroscience S-603, Baylor College of Medicine, Houston, Texas 77030, USA

FOR many types of behaviours, it is necessary to monitor the position or movement of objects that are temporarily occluded. The primate posterior parietal cortex contains neurons that are active during visual guidance tasks: in some cases, even if the visual target disappears transiently^{1,2}. It has been proposed that activity of this sort could be related to current or planned eye movements^{1,2}, but it might also provide a more generalized abstract representation of the spatial disposition of targets, even when they are not visible. We have recorded from monkey posterior parietal cortex while the animal viewed a visual stimulus that disappeared, and then, depending on experimental context, could be inferred to be either moving or stationary. During this temporary absence of the stimulus, about half of the neurons were found to be significantly more active on those trials in which the stimulus could be presumed to be moving rather than stationary. The activity was thus present in the absence of either sensory input or motor output, suggesting that it may indeed constitute a generalized representation of target motion.

Two rhesus monkeys were trained to maintain fixation within a 1° window while they viewed the trial types shown in Fig. 1a. On full vision trials, a stimulus 0.25° across appeared 12° from the fovea, and then, after a brief delay, moved towards the fixation spot. On occlusion trials the stimulus appeared as above, but then disappeared without moving and reappeared at 2° eccentricity moving towards the fixation spot, with the reappearance timed to be consistent with the stimulus moving continuously during the time that it was not visible. Full vision and occlusion trials were randomly interleaved during a block of trial presentations. Blocks of full vision and occlusion trials were interleaved with blocks of a third trial type, blink trials. On these trials, the stimulus appeared and disappeared exactly as in the occlusion trials, but then it reappeared in the same location, as if it had been stationary during the time it was not visible. Until the reappearance of the stimulus, the visual stimulation was identical on the occlusion and blink trials, the only difference being that the animal presumably inferred that the stimulus was moving during the time it was invisible on occlusion trials and inferred that the stimulus was stationary during the same period on blink trials.

Received 26 August; accepted 7 December 1994.

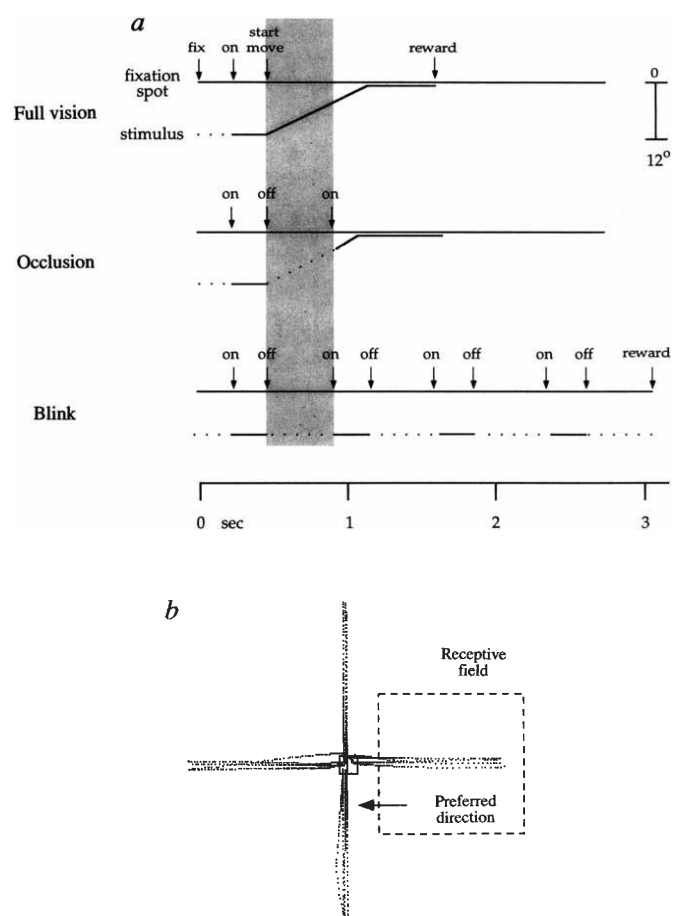
1. Ahissar, E. et al. *Science* **257**, 1412–1415 (1992).
2. Gray, C. M., Engel, A. K., König, P. & Singer, W. *Vis. Neurosci.* **8**, 337–347 (1992).
3. Eckhorn, R. et al. *Biol. Cybernet.* **60**, 121–130 (1988).
4. Vaadia, E. & Aertsen, A. in *Information Processing in the Cortex: Experiments and Theory* (eds Aertsen, A. & Braitenberg, V.) 81–102 (Springer, Heidelberg, 1992).
5. Murthy, V. N., Aoki, F. & Fetz, E. E. in *Oscillatory Event Related Brain Dynamics* (eds Pantev, C., Elbert, Th. & Lütkenhöner, B.) (Plenum, London, in the press).
6. Barlow, H. B. *Perception* **1**, 371–394 (1972).
7. Barlow, H. B. in *Information Processing in the Cortex: Experiments and Theory* (eds Aertsen, A. & Braitenberg, V.) 169–174 (Springer, Heidelberg, 1992).
8. Newsome, W. T., Britten, K. H. & Movshon, J. A. *Nature* **341**, 52–54 (1989).
9. Aertsen, A. M., Gerstein, G. L., Habib, M. K. & Palm, G. *J. Neurophysiol.* **61**, 900–917 (1989).
10. Toyama, K., Kimura, M. & Tanaka, K. *J. Neurophysiol.* **46**, 191–201 (1981).

We recorded responses to these stimuli from 106 neurons in the posterior parietal cortex (PP), in the lateral bank of the intraparietal sulcus and the fundus and anterior bank of the superior temporal sulcus. Figure 2 shows the average responses of two PP units to the three different trial types. The data shown are only for the preferred trajectory in each case. The histograms are synchronized (long vertical line) on the start of movement for the full vision trials, or on the disappearance of the stimulus for the occlusion and blink trials. On the occlusion and blink histograms, the small vertical line below the x-axis indicates the reappearance of the stimulus. Thus, all the activity between the synchronization and reappearance marks was evoked with the stimulus not visible (corresponding to the shaded area in Fig. 1a). Both cells responded vigorously to the moving stimulus on the full vision trials. There was also appreciable activity following disappearance of the stimulus on the occlusion trials, but not on the blink trials. This activity on the occlusion trials was not a simple sensory 'off' response because the visual stimulus was identical on the blink trials. Of the 106 cells examined, 48 (45%) showed a significantly different firing rate (averaged over the time that the stimulus was not visible) on the occlusion trials versus the blink trials (two-tailed *t*-test; $P < 0.05$). All 48 were more active on the occlusion trials. Figure 3a shows the distribution among all the units of a modulation index (MI) obtained by dividing the mean response of the occlusion trials by that of the blink trials, with the responses averaged as described above.

FIG. 1 Visual stimuli. *a*, Schematic of the three trial types. Time is shown on the horizontal axis; distance of the stimulus from the fixation spot on the vertical axis. Solid lines indicate that the stimulus is visible, dashed lines that it is not visible. Spike data were analysed over the shaded region—a time interval corresponding to, on the full vision trials, the start of movement until the stimulus was within 2° of the fixation spot or, on the other two trial types, the disappearance of the stimulus until its reappearance. On blink trials, the disappearance and reappearance of the stimulus was repeated four times to enhance the percept that the stimulus did not move. On full vision and occlusion trials, the animal received a juice reward after the stimulus reached the fixation spot; on blink trials the reward was delivered after completion of the fourth blink cycle. For comparison with the occlusion trials, data from blink trials were only examined during the first cycle, but the cells' responses generally did not change, or otherwise they declined on the later cycles. *b*, Typical stimulus trajectories, sampled at 75 Hz. While searching for units, trajectories were chosen randomly from trial to trial. Once a unit was isolated, its receptive-field size and directional preference were determined and then only four task trajectories were presented, spaced at 90° intervals, with the angle adjusted (while maintaining the 90° spacing) such that one traversed the receptive field of the unit, close to its preferred direction. As the trajectories were always centripetal, we necessarily selected units that had a major component of their preferred direction toward the fovea. During data collection, one block consisted of full vision and occlusion trials, randomly interleaved with respect to trial type and trajectory. The block continued until the animal had successfully completed two trials of each type and trajectory, for a total of 16 trials. The next block consisted solely of blink trials, again randomly interleaved with respect to the four initial locations, each repeated twice for a total of 8 trials. The timing of the disappearance and reappearance of the stimulus of each blink trial was identical to the timing of one of the corresponding occlusion trials of the previous block. This two-block alternation was then repeated 5–6 times for each unit, for a total of 10–12 repetitions of each trial type for each starting location. The trajectories used on full vision and occlusion trials were actually generated and stored during a previous block of trials in which the animal guided the stimulus to the fixation spot using a joystick. This accounts for the somewhat irregular trajectories exemplified in *b*. These movement trials will not be discussed; all the data herein were collected with the animals passively viewing as the stored stimulus templates were replayed. The start of each trial was indicated to the animal by a brief tone, with the pitch lower on the blink block so that the animal could distinguish the block type. Apart from this tone, the animal had no other way of distinguishing an occlusion trial from a blink trial on the first trial of a new block, because the visual

The indices of most cells were greater than unity; the median MI was 1.38. Population histograms among the 48 PP units that showed a significant difference between the two trial types ($O > B$ units) are shown in Fig. 3b. The fourth column is a difference histogram between the occlusion and blink trials. This difference in activity was sustained throughout the time that the stimulus was not visible.

A trivial explanation for the different responses between the occlusion and blink trials is that as the two were necessarily presented in separate blocks of trials, the animals' general state of arousal could have been lower during the blink blocks. Were this the case, it would be expected that the responses might differ throughout the trial, and not just subsequent to the disappearance of the stimulus. However, for the visual on-transient evoked by the initial onset of the stimulus prior to its disappearance, there was no significant difference between the two trial types among the 48 $O > B$ units (activity averaged over the 125-ms period following stimulus onset, a period before the beginning of the peristimulus-time histograms (PSTHs) in Figs 2 and 3; two-tailed paired *t*-test; $P > 0.20$; median MI, 1.05. For comparison, $P < 0.001$ for the same two-tailed paired *t*-test for the period following the disappearance of the stimulus). Thus the difference seen between occlusion and blink trials was not due to general differences in arousal between the two trial types. Another difference between the two trial types could be activity related to expectation of reward, because reward delivery occurred later



stimulus was identical until the reappearance of the stimulus. But any such confusion should have only diluted differences in response between the two trial types. Stimuli were presented on a video monitor refreshed at 75 Hz, placed 57 cm from the animal.

on the blink trials than on occlusion trials. A subset of the $O > B$ units (28/48) had bilateral receptive fields. For these units, the null-trajectory trials (for example, those trajectories moving from left to right in Fig. 1b) also traversed the receptive field. Responses from these cells (Fig. 4) showed no difference between occlusion and blink trials following the disappearance of the stimulus for these null-trajectory trials (Fig. 4; two-tailed paired t -test; $P > 0.50$; median MI, 0.91), suggesting that the difference on preferred direction trials could not be attributed to non-spatial factors such as anticipation of reward.

In addition to the sustained differences following the disappearance of the stimulus, there was a statistically significant increase in the signal before the disappearance of the stimulus for the preferred directions and a decrease for the null direction (averaged activity compared between occlusion or full vision versus blink trials for the period from 120 ms to 20 ms before the disappearance; one-tailed paired t -tests; $P < 0.001$). As the animals knew that the stimulus would reliably disappear a few hundred milliseconds after it appeared, the directionality of this early activity suggests that it may be related to the animals' expectation of the impending direction of motion.

In summary, a large proportion of neurons in monkey PP were more active following the disappearance of a visual stimulus on trials in which the animal could have inferred that the invisible stimulus was moving rather than stationary. The difference in activity was not immediately related to sensory input, because the stimulus was identical between the two trial types. It has long been appreciated that the activity of PP neurons can be influenced by extraretinal as well as visual sources^{3,4}; several studies have even reported responses in PP in the complete absence of visual stimuli during eye-movement tasks^{1,2,5,6}. As these tasks involve motor actions, the activity could be related

to the intention or execution of the movement. In the present study, however, the animals had no behavioural requirement beyond maintaining fixation, and never made eye movements to the stimulus. Thus the difference in activity between the occlusion and blink trials could not be related in a straightforward way to either sensory input or motor output. The only difference between the two trial types would seem to be the animals' inference that the invisible stimulus were either moving or stationary. The difference in activity could be a neuronal correlate of such inference. Signals of this sort might contribute to the extraretinal responses observed in visual guidance tasks such as smooth pursuit; for example, the activity in cells in the medial superior temporal area that persists when the pursuit target is transiently

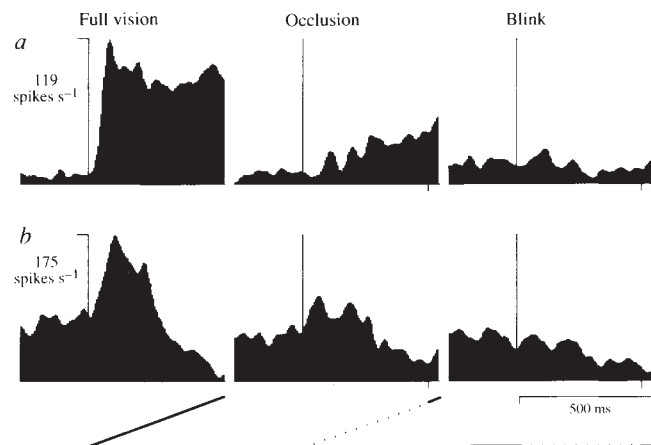


FIG. 2 Responses for preferred-trajectory trials from two PP units. *a*, Peristimulus-time histograms (PSTH) from one cell for the three trial types. *b*, Data from a different cell. Histograms were smoothed by convolving with a gaussian with a standard deviation of 5 ms. Each PSTH was compiled from 12 trials, synchronized (long vertical line) on the start of movement for the full vision trials, or on the disappearance of the stimulus on the occlusion and blink trials. The short vertical line below the x-axis on the occlusion and blink trials represents the earliest reappearance of the stimulus among the 12 trials used to compile the PSTH (as the timing among trials was not uniform). The traces at the bottom show the timing of the disappearance of the stimulus, as described for Fig. 1a. Single-unit recordings were made using tungsten electrodes (Microprobe Inc.) or glass-insulated Pt-Ir electrodes, introduced into cortex through a guide tube⁷. All surgery was performed aseptically and under full anaesthesia in accordance with Baylor College of Medicine and USDA regulations. The locations of recording sites were verified histologically in one animal; the second animal is still under study.

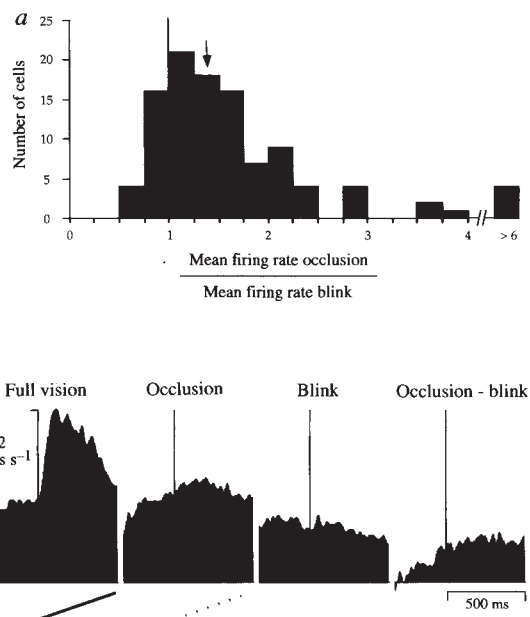


FIG. 3 Population responses. *a*, Distribution among all 106 units of the value of the mean rate of firing during occlusion trials divided by the mean rate during blink trials. Mean rates were calculated from the disappearance of the stimulus until its reappearance. Baseline activity was not subtracted in computing the ratio; this would have increased the absolute value of the modulation indices. The vertical line demarcates a ratio of one, and the arrow indicates the median value for the distribution. *b*, Population PSTHs compiled by calculating the average firing rate for each bin across the 48 PP neurons that showed a significant difference in mean firing rate between occlusion and blink trials. PSTHs were synchronized and smoothed as described for Fig. 2. The fourth column is a difference PSTH calculated by a bin-by-bin subtraction of the blink PSTH from the occlusion PSTH. A population PSTH was also compiled in identical fashion among 18 neurons recorded from area MT. The mean response on occlusion trials was not significantly greater than that on blink trials, averaged over the period following the disappearance of the stimulus (one-tailed paired t -test; $P > 0.50$; median MI, 0.90). For all cells, eye position was monitored using a scleral search coil system^{8,9} and stored at 200 Hz. An analysis was then performed for every preferred trajectory trial among the 48 $O > B$ units (~600 trials of each type) to test whether eye movements differed in a systematic fashion between the occlusion and blink trials during the period that the stimulus was not visible. For every pair of successive eye-position samples, the component of the instantaneous eye velocity parallel to the preferred stimulus trajectory was determined. The distribution of these velocities (~60,000 for each trial type) was then compiled and compared between occlusion and blink trials. No significant difference was found (Mann-Whitney test; $P > 0.05$. Median velocities were very close to zero in both cases: -0.0072° per second for occlusion trials and 0.0007° per second for blink trials).

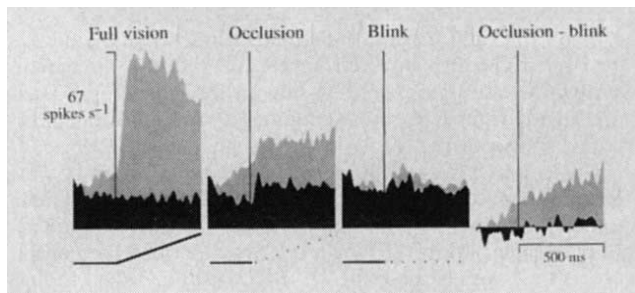


FIG. 4 Comparison of preferred and null-trajectory trials for the 28 $O > B$ units with bilateral receptive fields. Cells were defined to have bilateral receptive fields if their response to the onset of the stimulus was not greater for the stimulus corresponding to the preferred direction compared to the null direction (activity averaged over the 125-ms period following stimulus onset; one-tailed t -test; $P > 0.05$). Twenty-eight of the 48 $O > B$ cells met this criterion. PSTHs were synchronized and smoothed as described for Figs 2 and 3b. Preferred-direction responses are shown in grey, null-direction responses in black.

extinguished or stabilized on the retina¹. This activity might be related to the animal's presumption that a target is moving through space, rather than, or in addition to, motor-related signals such as efference copy from the motor system. □

Received 23 August; accepted 12 December 1994.

1. Newsome, W. T., Wurtz, R. H. & Komatsu, H. *J. Neurophysiol.* **60**, 604–619 (1988).
2. Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W. & Fogassi, L. *J. Neurosci.* **10**, 1176–1196 (1990).
3. Mountcastle, V. B., Lynch, J. C., Georgopoulos, A. P., Sakata, S. & Acuna, C. *J. Neurophysiol.* **38**, 871–908 (1975).
4. Robinson, D. A., Goldberg, M. E. & Stanton, G. B. *J. Neurophysiol.* **41**, 910–932 (1978).
5. Barash, S., Bracewell, R. M., Fogassi, L., Gnadt, J. W. & Andersen, R. A. *J. Neurophysiol.* **66**, 1095–1108 (1991).
6. Barash, S., Bracewell, R. M., Fogassi, L., Gnadt, J. W. & Andersen, R. A. *J. Neurophysiol.* **66**, 1109–1124 (1991).
7. Crist, C. F., Yamasaki, D. S. G., Komatsu, H. & Wurtz, R. H. *J. Neurosci. Meth.* **26**, 117–122 (1988).
8. Robinson, D. A. *IEEE Trans. biomed. Engng* **10**, 137–145 (1963).
9. Judge, S. J., Richmond, B. J. & Chu, F. C. *Vision Res.* **20**, 535–538 (1980).

ACKNOWLEDGEMENTS. We thank L. Colquhoun, J. Gibson, D. Leopold, N. Logothetis, D. Sheinberg and S. Treue for comments on the manuscript. This research was supported by awards from the National Eye Institute and the McKnight Foundation. J.A.A. was supported by a fellowship from the NIH.

Seeing the forest but only half the trees?

John C. Marshall & Peter W. Halligan*

Neuropsychology Unit, University Department of Clinical Neurology,
The Radcliffe Infirmary, Oxford OX2 6HE, UK

* Rivermead Rehabilitation Centre, Oxford OX1 4XD, UK

In visuo-spatial neglect after right hemisphere damage, patients fail overtly to notice and respond to stimuli on the side of space contralateral to lesion^{1,2}. Nonetheless, these neglected stimuli may covertly influence their performance in other tasks less direct than overt detection or identification¹. We report here a new type of dissociation between two forms of conscious perceptual awareness in a patient with left neglect. J.R. was shown hierarchical drawings in which a larger (global) form, such as a geometric figure or an alphabetic letter, is composed of smaller (local) forms (dots, circles, or letters). She gave accurate verbal reports of the global structure of these stimuli (Navon figures³), yet when required to cross out the smaller subfigures, she only cancelled those on the right of each global figure. Conscious perception of the whole does not automatically lead to visual awareness of all the parts thereof.

J.R., a 59-year-old right-handed woman, initially presented in 1993 with a right hemiplegia from which she rapidly made a full recovery and returned to work; a computed tomography (CT) scan showed a small lacunar infarct of the left thalamic region. A year later, she suffered a right hemisphere stroke, with left hemiplegia, left visual extinction, and gross left visuo-spatial neglect. The weakness improved in the left leg and to a lesser extent in the left arm, but the other symptoms persisted over three months after the second stroke. Florid left neglect was repeatedly found on three diagnostic tasks: presented with random arrays of simple figures to cancel, J.R. crossed out the stimuli on the right-hand side of the page, neglecting all those on the left. Requested to bisect horizontal lines, she placed her transections significantly to the right of centre. When copying complex line drawings, J.R. made relatively accurate representations of the right side of figures but omitted the left sides without acknowledging that the copies were incomplete. Drawing from memory was considerably better. CT scan showed infarction of frontal and parietal cortex with probable subcort-

ical involvement. The lacunar infarct from the earlier episode was still visible.

The following observations were made in the second and third months after the second stroke. J.R. was presented with Navon figures of the types illustrated in Fig. 1. Each stimulus, black on a sheet of white A4 paper, was placed on the desk top immediately in front of her. In free vision, and with no time constraint, J.R. was requested to (1) describe what she saw on the page; (2) cancel all the smaller subfigures (with a pen held in the right hand); (3) put down the pen when all components had been cancelled and describe again what she had seen. The figures varied in overall size and in the spacing between their constituents; some global figures were simple geometric forms, others were letters; the constituent (local) figures were dots (1 mm), circles (15 mm in diameter), or letters (5 mm × 5 mm).

J.R. invariably gave a correct verbal description of each locally consistent stimulus. Figure 1b (180 × 180 mm), for example, was reported as a large circle of small As; Fig. 1c (110 × 80 mm) was reported as a large E made of smaller Hs. Despite these accurate reports (before and after cancellation), J.R. never crossed out the subfigures on the left of the overall configuration. The same failure was seen when the cancellations were made in invisible ink, and is thus not due to attentional capture by the additional marks made by visible ink. Our dataset comprises over fifty observations of the type shown in Fig. 1; there were no counter-examples under the testing conditions described.

A number of shallow explanations can be ruled out. The conjecture that J.R. only 'saw' the right half of each square or circular configuration, but 'completed' the global percept by reporting the simplest closed figure compatible with what was seen⁴ is false: when shown 'full' figures (Navon squares and circles) randomly interspersed with 'half' figures, J.R. reported only the squares and circles as squares and circles ($n = 10$), while correctly describing the left-truncated figures as half-squares or half-circles ($n = 10$). By contrast, when shown circles and squares in which the (lateral) halves were composed respectively of Es and Fs (or Cs and Xs), reliable local completion occurred. All global stimuli were correctly reported as circles or squares, but J.R. claimed that the (correctly named) local elements on the right extended throughout each global figure ($n = 10$). When the (lateral) halves of a circle were respectively composed of blue and red dots, the colour of the right half was reported for all the dots of the circle ($n = 10$).

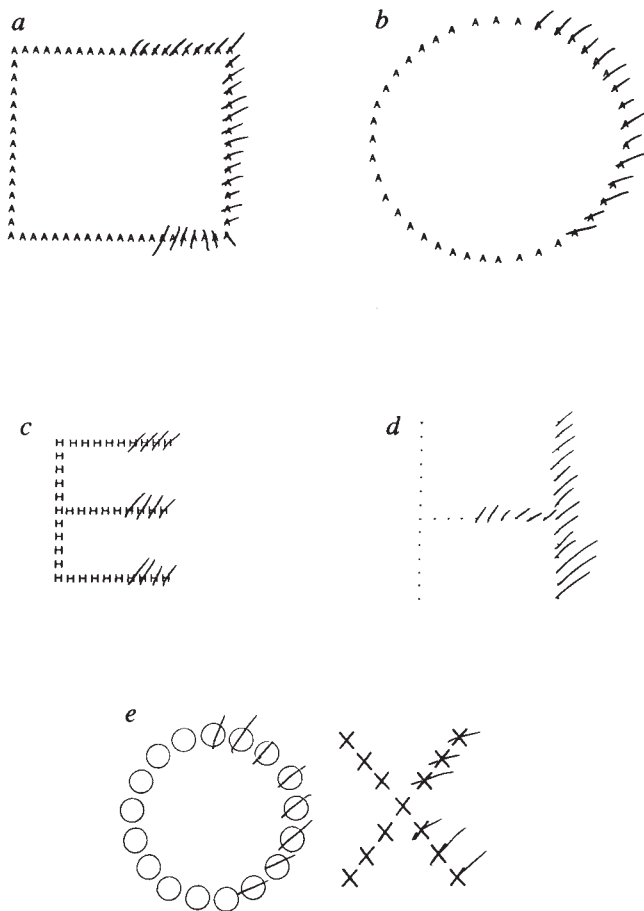


FIG. 1 J.R.'s attempts to cross out (cancel) the local components of global forms after correctly naming both aspects of each figure.

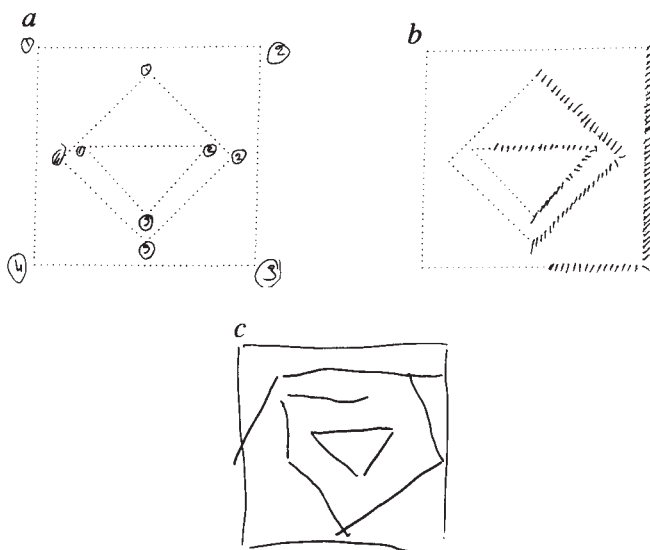


FIG. 2 J.R.'s marking of the corners of a complex pattern (a); her attempts to cancel all the local components (b); her copy of the stimulus display from memory (c).

The hypotheses that J.R. became fatigued after making cancellations on the right or that she cannot (under visual guidance) extend her right arm into left space⁵ are likewise inadequate: given two Navon figures side-by-side on the page (Fig. 1e), she cancelled the right half of each figure ($n=4$). A further falsification is shown in Fig. 2.

J.R. was presented with a hierarchical configuration in which a large square of dots (183×183 mm) contained a smaller diamond of dots (100×100 mm), which in turn contained a smaller triangle of dots (80×50 mm). She described this complex figure as we have just described it. J.R. was then asked to mark the corners of the configurations she had described. The result is shown in Fig. 2a; the square was numbered first, then the diamond and then the triangle. She was next given another identical stimulus which she again described accurately (and stated that it was identical to the first). We now requested that she cancel all the dots that she saw (Fig. 2b). The stimulus was then removed and J.R. was asked to draw from memory (with lines, not dots) the configuration she had just seen (Fig. 2c). Her recall of the display is intact.

J.R.'s neglect thus manifests when focal attention should be directed sequentially to components of a display. If her impairment is visuo-spatial (rather than purely spatial), J.R. should succeed better with endogenously driven than with comparable exogenously driven tasks: some patients with left neglect produce better drawings of common objects with their eyes closed than with their eyes open⁶. J.R. is relatively good at spontaneous drawing with her eyes open, but she improves on cancellation tasks with her eyes closed. We asked her to draw (continuous) squares and circles with her eyes open, to imagine that these drawings were composed of small dots, and to cancel them; the task was then repeated with J.R. blindfolded ($n=4$). In relevant respects, her performance was always better without vision. Figure 3a, drawn with the eyes open, shows left neglect; Fig. 3b, drawn blindfolded, does not.

This contrast between exogenously and (predominantly) endogenously driven tasks persisted even when J.R. had her eyes open. J.R. was again requested to cancel a circle of dots; left neglect is manifest (Fig. 4a). When the stimulus sheet was replaced by a blank sheet on which J.R. was asked to reproduce the pattern of dots that she had just seen, a full circle of dots was drawn (Fig. 4b). Her performance in the latter case is endogenously driven by the representation of the full circle of dots that she recalled; the specific demands upon focal attention that are required by the visual search task (cancellation) are reduced. Even when J.R. herself drew the circle of dots and then immediately cancelled them, left neglect was still evident on cancellation (Fig. 4c).

Each (normal) cerebral hemisphere orients attention primarily to contralateral space^{7,8}. The left hemisphere is also specialized for focal attention and the right hemisphere for more global processing⁹. Our results show that after right hemisphere damage the global perception of well structured objects remains possible

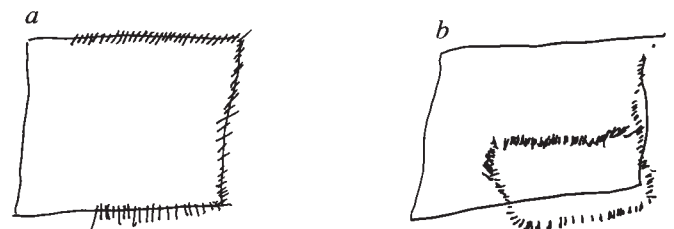


FIG. 3 J.R.'s attempts to draw a square and cancel the imaginary components thereof: with (a) and without (b) vision.

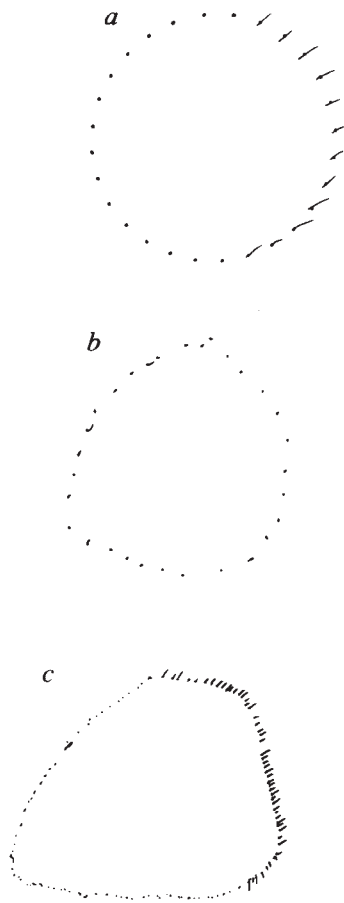


FIG. 4 J.R.'s attempts to cancel the local components of a predrawn figure (a); to reproduce the figure (b); and to draw the figure and then cancel its local components (c).

even when focal attention cannot be automatically directed to local components on the left of a global figure. J.R.'s neglect is thus neither purely perceptual nor purely motor; it rather arises from an input-deficit to a 'premotor' mechanism¹⁰ committed to the exogenous control of focal attention. Lesions that 'disrupt the cross-talk between inferior posterior areas' can reduce or eliminate interactions between global and local processing⁹. Whether this disruption in J.R. is due solely to her right hemisphere lesion (or also implicates her left thalamic infarct) is unclear. One further problem remains: why is J.R. unaware of the discrepancy between her global percepts and local actions? The process of local completion can mask perceptual incongruity but cannot then control appropriate manual output. J.R. can perceive the whole forest but cannot use that percept to search for and cut down the trees on the left thereof. □

Received 6 September; accepted 12 December 1994.

1. Marshall, J. C. & Halligan, P. W. *Nature* **336**, 766–787 (1988).
2. Halligan, P. W. & Marshall, J. C. *Cogn. Neuropsychol.* **11**, 167–206 (1994).
3. Navon, D. *Cogn. Psychol.* **9**, 353–383 (1977).
4. Sergent, J. *Brain* **111**, 347–373 (1988).
5. Bottini, G., Sterzi, R. & Vallar, G. *J. Neurol. Neurosurg. Psychiatr.* **55**, 562–565 (1992).
6. Halligan, P. W. & Marshall, J. C. *Cortex* **30**, 685–694 (1994).
7. Kinsbourne, M. *Trans. Am. Neurol. Assoc.* **95**, 143–145 (1970).
8. Ladavas, E., Del Pesce, M., Mangun, G. R. & Gazzaniga, M. S. *Cogn. Neuropsychol.* **11**, 57–74 (1994).
9. Robertson, L. C. in *Cognitive Neuropsychology in Clinical Practice* (ed. Margolin, D. I.) 70–95 (Oxford University Press, New York, 1992).
10. Rizzolatti, G. & Berti, A. *Rev. Neurol.* **146**, 626–634 (1990).

ACKNOWLEDGEMENTS. This work was supported by the Medical Research Council.

Alzheimer-type neuropathology in transgenic mice overexpressing V717F β -amyloid precursor protein

Dora Games*, David Adams††, Ree Alessandrini†, Robin Barbour*, Patricia Berthelette††, Catherine Blackwell††, Tony Carr*, James Clemens§, Thomas Donaldson††, Frances Gillespie††, Terry Guido*, Stephanie Hagopian††, Kelly Johnson-Wood*, Karen Khan*, Mike Lee*, Paul Leibowitz††, Ivan Lieberburg*, Sheila Little§, Eliezer Masliah||, Lisa McConlogue*, Martin Montoya-Zavala††, Lennart Mucke*, Lisa Paganini*, Elizabeth Penniman†, Mike Power*, Dale Schenk*, Peter Seubert*, Ben Snyder†, Ferdie Soriano*, Hua Tan*, James Vitale††, Sam Wadsworth††, Ben Wolozin** & Jun Zhao*

* Athena Neurosciences, Inc., 800 Gateway Boulevard, South San Francisco, California 94080, USA

† Exemplar Corporation, One Innovation Drive, Worcester, Massachusetts 01605, USA

§ Lilly Research Laboratories, Indianapolis, Indiana 46285, USA

☆ The Scripps Research Institute, Department of Neuropharmacology, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

|| Department of Neurosciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

** Laboratory of Clinical Science, National Institute of Mental Health, 9000 Rockville Pike, 1013D41, Bethesda, Maryland 20892, USA

ALZHEIMER'S disease (AD) is the most common cause of progressive intellectual failure in aged humans. AD brains contain numerous amyloid plaques surrounded by dystrophic neurites, and show profound synaptic loss, neurofibrillary tangle formation and gliosis. The amyloid plaques are composed of amyloid β -peptide (A β), a 40–42-amino-acid fragment of the β -amyloid precursor protein (APP)¹. A primary pathogenic role for APP/A β is suggested by missense mutations in APP that are tightly linked to autosomal dominant forms of AD^{2,3}. A major obstacle to elucidating and treating AD has been the lack of an animal model. Animals transgenic for APP have previously failed to show extensive AD-type neuropathology^{4–10}, but we now report the production of transgenic mice that express high levels of human mutant APP (with valine at residue 717 substituted by phenylalanine) and which progressively develop many of the pathological hallmarks of AD, including numerous extracellular thioflavin S-positive A β deposits, neuritic plaques, synaptic loss, astrogliosis and microgliosis. These mice support a primary role for APP/A β in the genesis of AD and could provide a preclinical model for testing therapeutic drugs.

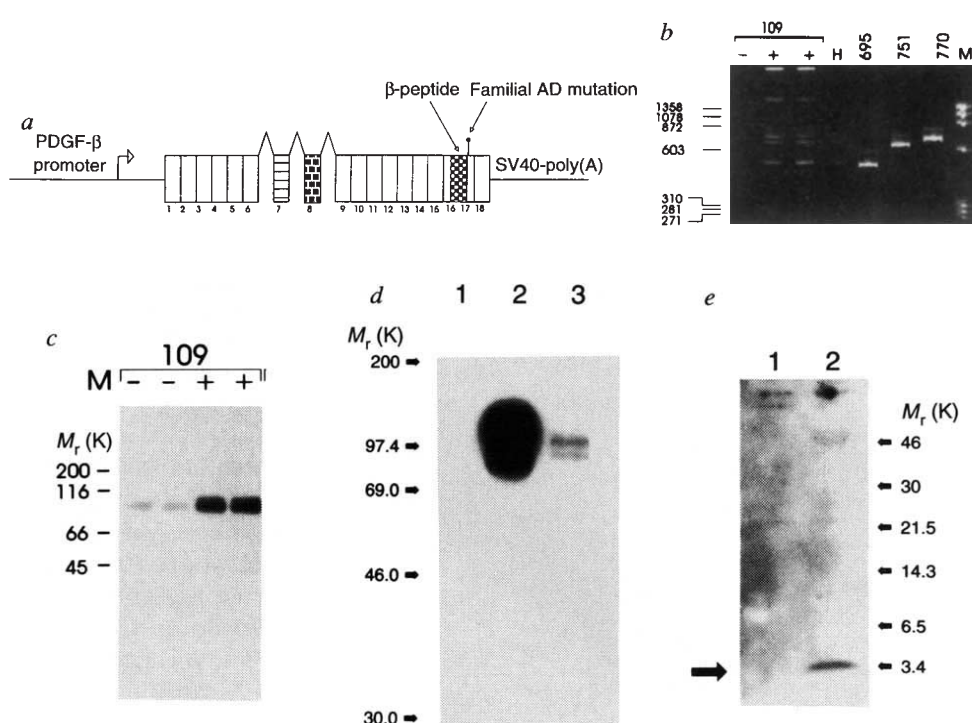
Transgenic mice were generated using a platelet-derived growth factor (PDGF)- β promoter¹¹ driving a human APP (hAPP) minigene encoding the APP_{717V→F} mutation associated with familial AD¹² (PD-APP; Fig. 1a). The construct contained APP introns 6–8, allowing alternative splicing of exons 7 and 8.

† Present addresses: Department of Biology and Biotechnology, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA (D.A.); Genzyme Corporation, One Mountain Road, Framingham, MA 01701, USA (P.B.); University of Massachusetts, Department of Veterinary and Animal Sciences, Amherst, MA 01003, USA (C.B.); Molecular Therapeutics Inc., 400 Morgan Lane, West Haven, CT 06516, USA (T.D.); Transkaryotic Therapies Inc., 195 Albany Street, Cambridge, MA 02139, USA (F.G.); Joint Program in Neonatology, Children's Hospital, Enders, Rm 950, 300 Longwood Avenue, Boston, MA 02115, USA (S.H.); Innovir Laboratories Inc., 510 East 73 Street, New York, NY 10022, USA (P.L.); Alton Inc., 170 Williams Drive, Ramsey, NJ 07446, USA (M.M.-Z.); Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA (J.V.).

¶ To whom correspondence should be addressed.

Fig. 1 analysis of app expression in brain tissue of pd-app transgenic mice. **a**, Map of the APP construct (PD-APP) minigene used to generate transgenic mice (not to scale). The construct contains the PDGF β -chain promoter, full-length hAPP cDNA encoding the Val-to-Phe mutation at codon 717, and the inclusion of genomic sequences for hAPP introns 6–8. **b**, RT-PCR analysis demonstrates the presence of transcripts encoding the 695, 751 and 770 isoforms of hAPP in transgenic animal brains (+) but not in brains from non-transgenic littermates (–). Control reactions using human brain RNA (H) as well as cDNA clones encoding hAPP 695, 751 or 770 are also shown. M, Markers (calibration in bp). **c**, Immunoblot analysis of total APP expression (human and mouse) in transgenic mouse (+) and control littermate (–) brain tissue using a C-terminal APP antibody, $\alpha 6$. **d**, Human-specific APP expression in brain tissue from a 6-month-old non-transgenic littermate (lane 1), transgenic mouse (lane 2) or human AD cortex (lane 3) using immunoblotting with the human-specific APP antibody 8E5. **e**, Immunoprecipitation and immunoblot analysis of $A\beta$ from the brain tissue of a 9-month-old non-transgenic littermate (lane 1) and a 9-month-old transgenic mouse (lane 2).

METHODS. The PDGF β -chain 5' flanking sequence included 1.3 kb upstream of the transcription initiation site and ~70 bp of 5' untranslated region, ending at the *AurII* site¹⁰. The β -APP sequences were derived from human β -APP cDNA starting with the *NruI* site in exon 1 through to the end of exon 6 and from the beginning of exon 9 through to exon 18, including the 3' untranslated sequence ending at the *SphI* site, and from human genomic sequences from exons 6–9, including all intervening introns. To introduce the familial AD mutation at position 717, Val was mutated to Phe (G to T, using the P-Select mutagenesis kit; Promega). The late SV40 polyadenylation signal was provided by the 240-bp *BamHI* to *BclI* fragment. Plasmid sequences (pUC) were removed by *SacI* and *NotI* digestion before microinjection. Transgenic mice were generated using standard techniques¹⁹, except that PD-APP DNA was microinjected into the embryos at the two-cell stage. Seven founder mice were generated and line 109 was used for extensive analysis. RNA was isolated from brain tissue as described²⁰ and subjected to RT-PCR as described²¹ using human-specific APP primers (5'-CCGATGATGACGAGGACGAT-3', 5'-TGAACACGTGACGAGGCCGA-3') with



the following PCR conditions: 40 cycles of 1 min at 94 °C, 40 s at 60 °C, 50 s at 72 °C. The identities of the human APP RT-PCR bands from the transgenic mouse RNA were verified by subcloning and sequencing. Analysis of holo-APP involved brain homogenization in 10 volumes of PBS containing 0.5 mM EDTA, 10 $\mu\text{g ml}^{-1}$ leupeptin and 1 mM PMSF. Samples were spun at 12,000g for 10 min and the pellets resuspended in RIPA (150 mM NaCl, 50 mM Tris, pH 8.0, 20 mM EDTA, 1.0% deoxycholate, 1.0% Triton X-100, 0.1% SDS, 1 mM PMSF and 10 $\mu\text{g ml}^{-1}$ leupeptin (d)). Samples (each containing 30 μg total protein) were analysed by SDS-PAGE, transferred to Immobilon and reacted with either the holo-APP antibody, $\alpha 6$ (ref. 22), or 8E5 monoclonal antibody. 8E5 was prepared against a bacterial fusion protein encompassing hAPP residues 444–592 (ref. 22) and is human-specific showing essentially no crossreactivity against mouse APP (d, lane 1). For immunoblot analysis of $A\beta$ (e), a 9-month-old mouse brain was homogenized in 5 ml 6 M guanidine HCl, 50 mM Tris, pH 7.5. The homogenate was centrifuged at 100,000g for 15 min and the supernatant was dialysed against H_2O overnight adjusted to PBS with 1 mM PMSF and 25 $\mu\text{g ml}^{-1}$ leupeptin. This material was immunoprecipitated with antibody 266 resin, and immunoblotted with the human-specific $A\beta$ antibody, 6C6, as described¹³.

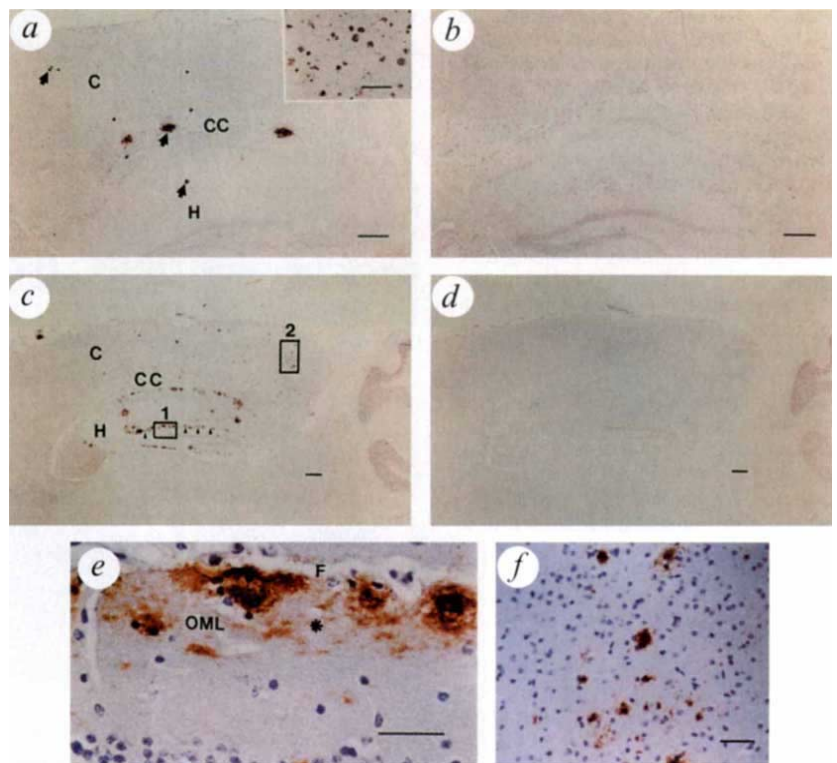
We used only heterozygous animals. Southern analysis of 104 from four generations showed that ~40 copies of the transgene were inserted at a single site and transmitted in a stable manner (data not shown). Human APP messenger RNA was produced in several tissues of the transgenic mouse, but at especially high levels in brain (not shown). RNase protection assays revealed ~18-fold more APP expression in the brains of line 109 animals than in most previously described lines expressing neuron-specific enolase (NSE)-promoter-driven APP transgenes (not shown; refs 4, 8–10). In addition, the three major splicing variants of hAPP mRNA (695, 751, 770)¹ were expressed in the transgenic mice, as evidenced by reverse-transcriptase polymerase chain reaction (RT-PCR) (Fig. 1b). Immunoblot analysis of brain homogenates using either a holo-APP polyclonal antibody or a human-specific APP monoclonal antibody revealed hAPP overexpression in the transgenic mouse at levels ≥ 10 -fold higher than either endogenous mouse APP levels (Fig. 1c, d) or those in AD brain (Fig. 1d). Using human-specific $A\beta$ antibodies, we isolated a 4K $A\beta$ -immunoreactive peptide from the brains of the transgenic animals, which corresponds to the relative molecular mass of $A\beta$ (Fig. 1e). Brain levels of $A\beta$ were at least 10-fold higher

in line-109 animals than in the previously described hAPP transgenic mice (not shown; ref. 9). Finally, embryonic day-16 cortical cell cultures from transgenic animals constitutively secreted human $A\beta$, including a substantial fraction of $A\beta$ 1–42 (5 ng ml^{-1} total $A\beta$; 0.7 ng ml^{-1} $A\beta$ 1–42), as detected in media by human-specific $A\beta$ enzyme-linked immunosorbent assays (not shown; refs 9 and 13). Thus, line-109 animals greatly overexpressed human APP mRNA, holo-APP and $A\beta$ in their brains.

Brains from 18 transgenic animals and 12 age-matched non-transgenic littermate controls (4–13 months old) representing three generations of the line-109 pedigree were extensively examined histopathologically. Between 4–6 months of age ($n=7$), no obvious pathology was detected; however, at ~6–9 months of age ($n=7$), transgenic animals began to exhibit deposits of human $A\beta$ in the hippocampus, corpus callosum and cerebral cortex, but not in other brain regions. These increased with age, and by eight months many deposits (30–200 μm) were seen (Fig. 2a). As the animals aged (≥ 9 months; $n=4$), the density of the plaques increased (Fig. 2c) until the $A\beta$ -staining pattern resembled that of AD (Fig. 2a, inset). Robust pathology was

FIG. 2 Demonstration of AD-like amyloid plaques in PD-APP transgenic mice. Human and mouse brain sections were labelled with antiserum R1280 generated against synthetic human A β 1–40 peptide. H, Hippocampus; C, cortex; CC, corpus callosum; OML, outer molecular layer of the dentate gyrus; F, hippocampal fissure. Coronal sections of the hippocampus and neocortex from *a*, an 8-month-old transgenic mouse containing A β deposits (arrows), and *b*, a non-transgenic littermate. Inset in *a*, human tissue from AD frontal cortex stained with R1280. Adjacent parasagittal sections from a 13-month-old transgenic mouse before (*c*) and after (*d*) preincubation of the antibody with synthetic A β 1–40 peptide. In *c*, the large increase in A β was confined to the cortex and hippocampus. Several regions of the hippocampus contained densely packed A β , including the terminal zone of the perforant pathway in the outer molecular layer of the dentate gyrus (arrowheads). Boxed areas 1 and 2 are shown at higher magnification in *e* and *f*, respectively. Scale bars in *a*–*d*, 200 μ m. In *e*, the outer molecular layer of the dentate gyrus contained areas of compacted and diffuse (asterisk) plaques. The edge of the granule cell layer is visible at the bottom. In *f*, a field of A β deposits in the occipital cortex. Scale bars in *e* and *f*, 40 μ m.

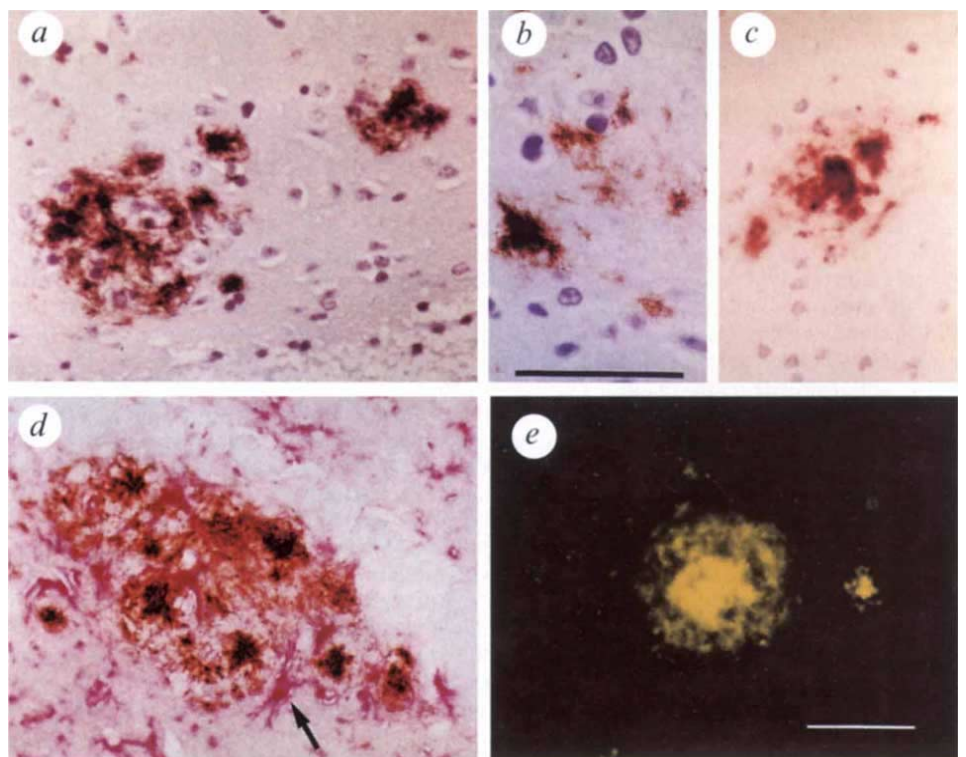
METHODS. Mouse brains were removed and placed in Trojanowski's fixative²³ for 48 h before paraffin embedding. 6- μ m coronal or parasagittal sections from transgenic and non-transgenic mice were placed adjacent to each other on poly-L-lysine coated slides. Sections were deparaffinized, rehydrated, and treated with 0.03% H₂O₂ for 30 min before overnight incubation at 4 °C with a 1:1,000 dilution of the A β antibody, R1280 (ref. 24). For absorption studies, synthetic human A β 1–40 peptide²⁵ in 10% aqueous dimethylsulphoxide was added to a final concentration of 7.0 μ M to the diluted antibody and incubated for 2 h at 37 °C. The diluent was applied to the sections and processed under the same



conditions as the standard antibody solution. Peroxidase rabbit IgG kit (Vector Labs) was then used as recommended, with 3,3'-diaminobenzidine (DAB) as the chromagen. Similarly fixed human AD brain was processed simultaneously under identical conditions.

FIG. 3 Morphological diversity of A β deposition in the PD-APP mouse brain. Roughly spherical (*a*), and wispy, irregular deposits (*b*), labelled with antibody 9204 (A β 1–5; ref. 26) specific for the free N terminus of A β . *c*, A β core and surround labelled with antibody 277-2, specific for the C terminus of A β 1–42. *d*, Astrocytic gliosis (arrow) associated with A β deposition was evident after double immunolabelling with antibodies to glial fibrillary acidic protein (GFAP, red) and human A β (A β 1–5; brown). *e*, A β deposits were also reactive with thioflavin S. A compacted A β core and 'halo' is evident in the large plaque. The fine background fluorescence represents autofluorescent lipofuscin granules. Scale bars, 50 μ m.

METHODS. For *a*–*c*, immunocytochemistry was performed as described for Fig. 2. Antibody 9204 (to A β 1–5) was used at a concentration of 7.0 μ g ml^{–1}. Antibody 277-2, specific for A β 1–42 (apparent affinity for A β 1–42 = 75 nM versus >10 μ M for A β 1–40 by radioimmunoassay competition), was prepared by immunizing New Zealand white rabbits with the peptide cysteine-aminoheptanoic acid–A β 33–42 conjugated to cyanized BSA ('Super Carriers'; Pierce) using a typical immunization protocol (500 μ g per injection). Specific antibodies were affinity-purified from serum against the immunogen immobilized on agarose beads. Before incubation with 277-2 (10 μ g ml^{–1}) sections were treated for 1–2 min with 80% formic acid. In *c*, the antibody 9204 was reacted using the peroxidase rabbit IgG kit (Vector Labs). The product was then visualized with DAB, and the sections were incubated overnight at 4 °C with a 1:500 dilution of polyclonal anti-GFAP (Sigma). The GFAP antibody was reacted using

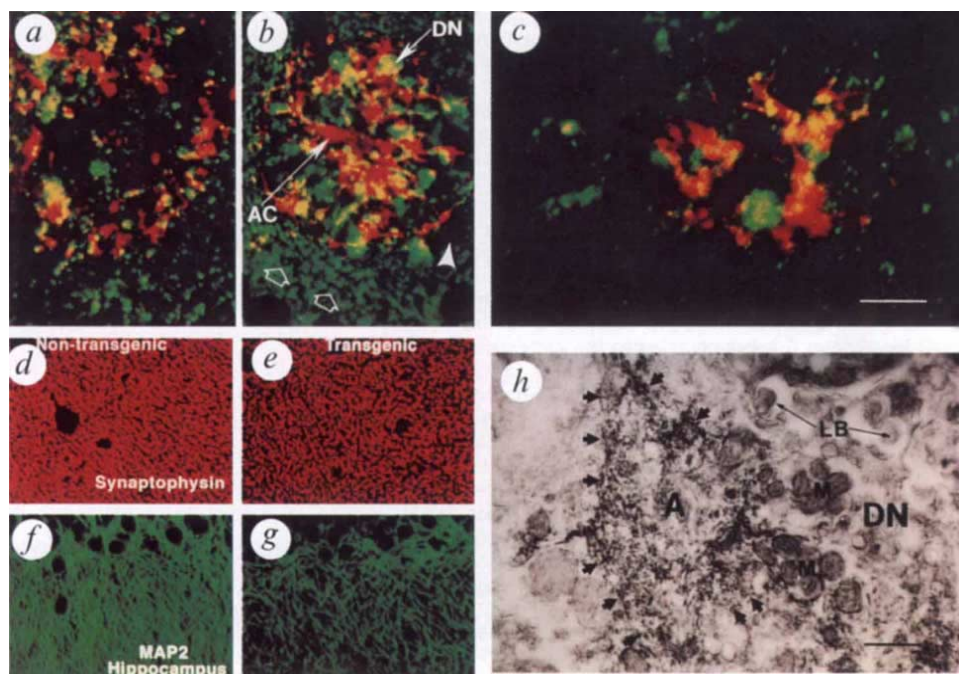


the alkaline phosphatase anti-rabbit IgG kit and alkaline phosphatase substrate kit I, (Vector Labs; used according to the manufacturer's recommendations). Sections were stained with thioflavin S using standard procedures²⁷ and viewed with ultraviolet light through an FITC filter of maximum wavelength 440 nm.

FIG. 4 Laser scanning confocal images of human and 8-month-old transgenic mouse brains, demonstrating the relationship of extracellular cortical A β deposit to dystrophic neurites and neuropil abnormalities, as well as the reduction of synaptic density and dendrites in the transgenic hippocampus. An A β deposit and adjacent neuropil in an 8-month-old transgenic mouse brain are also shown in an electron micrograph.

In both human (a) and mouse (b) brains, A β deposits (red) were associated with distorted neurites (DN) containing synaptophysin (green). Yellow signifies overlap of the two markers. b, A β immunoreactive plaque in a transgenic mouse brain containing a central amyloid core (AC); synaptic loss (arrowhead) and compression (open arrows) of the neuropil surrounding the amyloid deposit are also evident. c, A β deposit (red) in a transgenic mouse brain associated with morphologically abnormal hAPP-positive neurites (green). The magnification for a–g is indicated by the scale bar in c (20 μ m). Both synaptophysin (red) and microtubule-associated protein-2 (MAP-2; green) immunostaining were reduced throughout the molecular layer of the hippocampal dentate gyrus in the transgenic mouse (e, g) compared to the non-transgenic littermate (d, f). h, Immunoelectron micrograph of a transgenic mouse brain demonstrating extracellular A β deposition (A) decorated with the human-specific R1280 A β antiserum (outlined by arrows). A dystrophic neurite (DN) in the immediate vicinity of the A β deposit contained abundant large mitochondria (M) and laminar dense bodies (LB). Scale bar, 2 μ m.

METHODS. 40- μ m-thick vibratome sections were incubated overnight at 4 °C with the following antibodies: R1280 (1:1,000) in combination with polyclonal anti-synaptophysin in a and b (1:150; Dako) or 8E5 in c (7.0 μ g ml⁻¹; described in Fig. 1). For d–g, sections were incubated with anti-synaptophysin or monoclonal anti-MAP 2 (1:20, Boehringer-Mannheim), and reacted with a goat anti-rabbit biotinylated antibody (1:100) followed by a mixture of FITC-conjugated horse anti-mouse



IgG (1:75) and avidin D Texas red (1:100) (Vector Labs). The double-immunolabelled sections were viewed on a Zeiss Axiovert 35 microscope with attached laser confocal scanning system MRC 600 (Bio-Rad). The Texas red channel collected images of the R1280 (a, b, c) or synaptophysin (d, e) labelling, and the FITC channel collected synaptophysin (a, b), 8E5 (c), or MAP 2 (f, g) labelling. Optical z-sections 0.5 μ m in thickness were collected from each region; details of similar image processing and storage are described in ref. 28. For immunoelectron microscopy, mice were perfused with saline followed by 2.0% paraformaldehyde and 1.0% glutaraldehyde in cacodylate buffer. 40- μ m-thick vibratome sections were incubated with R1280, and reacted as in Fig. 2. Immunolabelled sections with A β deposits were then fixed in 1.0% ammonium tetroxide and embedded in epon/araldite before viewing ultrathin sections with a Jeol CX100 electron microscope²⁹.

also seen in another transgenic line generated from the PDAPP vector (line 35; data not shown). A β deposits of varying morphology clearly were evident as a result of using a variety of A β antibodies, including well characterized human-specific A β antibodies (Figs 2, 3) and antibodies specific for the free amino and carboxy termini of A β 1–42 (Fig. 3a–c). Serial sections demonstrated many plaques were positively stained with both of the latter antibodies. The forms of the A β deposition ranged from diffuse irregular types to compacted plaques with cores (Fig. 3). Non-transgenic littermates (Fig. 2b) showed none of these neuropathological changes. Immunostaining was fully absorbable with the relevant synthetic peptide (Fig. 2d), and was apparent using a variety of processing conditions, including fixation with paraformaldehyde and Trojanowski methods (Figs 2, 3). Many plaques were stained with thioflavin S (Fig. 3e), and some were also stained using the Bielschowsky silver method and were birefringent with Congo red (not shown), indicating the true amyloid nature of these deposits. Confirmation of the presence of extracellular A β was obtained using immunoelectron microscopy (Fig. 4h). The majority of plaques were intimately surrounded by GFAP-positive reactive astrocytes (Fig. 3d), similar to the gliosis found in AD plaques. The neocortices of the transgenic mice contained diffusely activated microglial cells, as defined by their amoeboid appearance and shortened processes (not shown). Preliminary attempts to identify neurofibrillary tangles with tau antibodies were negative, consistent with their well known absence in rodent tissues¹⁴. Nevertheless, clear

evidence for neuritic pathology was apparent using both conventional and confocal immunomicroscopy. Many A β plaques were closely associated with distorted neurites that could be detected with hAPP-specific antibodies (Fig. 4c) and with anti-synaptophysin antibodies (Fig. 4b), suggesting that these neurites were derived in part from axonal sprouts, as observed in the AD brain (Fig. 4a). The plaques compressed and distorted the surrounding neuropil (Fig. 4b), also as in the AD brain (Fig. 4a). Finally, synaptic and dendritic density were reduced in the molecular layer of the hippocampal dentate gyrus of the transgenic mice. This was evident by reduced immunostaining for the presynaptic marker synaptophysin (compare Fig. 4d and e) and the dendritic marker MAP-2 (compare Fig. 4f and g), as described in AD brain¹⁵.

Several transgenic rodent lines have been produced that express either the hAPP gene or hAPP complementary DNAs regulated by a variety of promoters^{4–10}. In particular, NSE-driven APP₇₅₁ transgenic mice^{9,10} have sparse A β deposits which are more typical of early AD and young Down's syndrome cases; in these mice, unlike ours, mature lesions such as frequent compacted plaques, neuritic dystrophy and extensive gliosis were very rare¹⁰. Our success in generating AD-like pathology consistently in these transgenic mice is probably due to the construct used and the high level of hAPP expression. The transgene contains a splicing cassette that permits expression of all three major hAPP isoforms. Expression is driven by the PDGF- β promoter, which is known to target expression preferentially to neurons

of the cortex, hippocampus, hypothalamus and cerebellum of transgenic animals¹¹. The familial AD mutation at residue 717¹² may be important as it partially shifts production of A β from the 40-amino-acid form to the more amyloidogenic 42-residue peptide known to predominate in plaques^{16,17}. Preparation of APP transgenic mice independently harbouring each of these features will be required to identify the essential component(s) that result in pathology.

The most notable feature of these transgenic mice is their Alzheimer-like neuropathology, which includes extracellular A β deposition, dystrophic neuritic components, gliosis and loss of synaptic density with regional specificity resembling that of AD. Based on the limited sampling to date, plaque density appears to increase with age in these transgenic mice, as it does in humans¹, implying a progressive A β deposition that exceeds its clearance, as also proposed for AD¹⁸. Our transgenic model provides strong new evidence for the primacy of APP expression and A β deposition in AD neuropathology and offers a means to test whether compounds that lower A β production and/or reduce its neurotoxicity *in vitro* can produce beneficial effects in an animal model prior to advancing such drugs into human trials. □

Received 17 October 1994; accepted 5 January 1995.

1. Selkoe, D. J. *A. Rev. Neurosci.* **17**, 489–517 (1994).
2. Hardy, J. *Clin. Geriatr. Med.* **10**, 239–247 (1994).
3. Mann, D. M. A. *et al. Neurodegeneration* **1**, 201–215 (1992).

4. Quon, D. *et al. Nature* **352**, 239–241 (1991).
5. Sanhu, F. A., Salim, M. & Zain, S. B. *J. Biol. Chem.* **266**, 21331–21334 (1991).
6. Lamb, B. T. *et al. Nature Genet.* **5**, 22–30 (1993).
7. Pearson, B. E. & Choi, T. K. *Proc. natn. acad. Sci. U.S.A.* **90**, 10578–10582 (1993).
8. Mucke, L. *et al. Brain Res.* **666**, 151–167 (1994).
9. McConlogue, L. *et al. Neurobiol. Aging* **15**, S12 (1994).
10. Higgins, L. S. *et al. Ann. Neurol.* **35**, 598–607 (1994).
11. Sasahara, M. *et al. Cell* **64**, 217–227 (1991).
12. Murrell, J. *et al. Science* **254**, 97–99 (1991).
13. Seubert, P. *et al. Nature* **359**, 325–327 (1992).
14. Cork, L. C. *et al. J. Neuropath. exp. Neurol.* **47**, 629–641 (1988).
15. Masliah, E. *et al. Am. J. Path.* **138**, 235–246 (1991).
16. Suzuki, N. *et al. Science* **264**, 1336–1340 (1994).
17. Roher, A. *et al. J. Biol. Chem.* **269**, 3072–3083 (1993).
18. Maggio, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 5462–5466 (1992).
19. Hogan, B., Constantini, F. & Lacey, E. in *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1986).
20. Chomazynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).
21. Wang, A. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9717–9721 (1989).
22. Oltersdorf, T. *et al. J. Biol. Chem.* **265**, 4492–4497 (1990).
23. Arai, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 2249–2253 (1990).
24. Tamaoka, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 1345–1349 (1992).
25. Games, D. *et al. Neurobiol. Aging* **13**, 569–576 (1992).
26. Saido, T. C. *et al. J. Biol. Chem.* **269**, 15253–15257 (1994).
27. Dickson, D. W. *et al. Acta neuropath.* **79**, 486–493 (1990).
28. Masliah, E. *et al. J. Neuropath. exp. Neurol.* **52**, 619–632 (1993).
29. Masliah, E. *et al. Acta neuropath.* **81**, 428–433 (1991).

ACKNOWLEDGEMENTS. We thank M. Mallory, A. Maya, N. Ge, E. Rockenstein, O. Stephenson and E. Johnstone for technical assistance; the animal care technicians involved in this project; J. Trojanowski for human brain sections; T. Saido and D. Selkoe for anti-A β antibodies 9204 and R1280, respectively; T. Collins and D. Goldgaber for the PDGF promoter and APP cDNA subclones, respectively; and D. Selkoe for helpful comments and advice.

Gating of the voltage-dependent chloride channel CIC-0 by the permeant anion

Michael Pusch, Uwe Ludewig, Annett Rehfeldt* & Thomas J. Jentsch†

Centre for Molecular Neurobiology (ZMNH), Hamburg University, Martinistrasse 52, D-20246 Hamburg, Germany

CHLORIDE channels of the CIC family are important for the control of membrane excitability^{1–3}, cell volume regulation^{4,5}, and possibly transepithelial transport^{6,7}. Although lacking the typical voltage-sensor found in cation channels^{8–10}, gating of CIC channels is clearly voltage-dependent. For the prototype *Torpedo* channel CIC-0 (refs 11–15) we now show that channel opening is strongly facilitated by external chloride. Other less permeable anions can substitute for chloride with less efficiency. CIC-0 conductance shows an anomalous mole fraction behaviour with Cl[−]/NO₃[−] mixtures, suggesting a multi-ion pore. Gating shows a similar anomalous behaviour, tightly linking permeation to gating. Eliminating a positive charge at the cytoplasmic end of domain D12 changes kinetics, concentration dependence and halide selectivity of gating, and alters pore properties such as ion selectivity, single-channel conductance and rectification. Taken together, our results strongly suggest that in these channels voltage-dependent gating is conferred by the permeating ion itself, acting as the gating charge.

The *Torpedo* electric organ Cl[−]-channel CIC-0 (ref. 12) has a 'slow' gate operating on both protochannels of the double-barrelled channel^{13,14} simultaneously, and a 'fast' gate acting on single protochannels^{13–15}. Both gates have opposite voltage-dependence, with the fast gate being opened by depolarization.

CIC-0 was expressed in *Xenopus* oocytes and the fast gate was studied in isolation (Fig. 1a). The dependence of the steady-state open probability, p_{open} , on the transmembrane voltage V can be described by the Boltzmann distribution $p_{\text{open}} = 1 / (1 + \exp(z_n e_0 (V_{1/2} - V) / kT))$ (where e_0 is the elementary charge, $V_{1/2}$ is the voltage of half-maximal activation, k is the Boltzmann constant, and T is the temperature) with a nominal gating charge $z_n \sim 1$ in agreement with earlier data^{13,14}. Consistent with single-channel measurements^{13,14}, gating kinetics indicate a two-state gating mechanism. Reducing extracellular Cl[−] concentration ($[\text{Cl}^-]_o$) shifts $p_{\text{open}}(V)$ to positive voltages without significantly changing its slope (the gating charge) (Fig. 1b). In contrast, intracellular chloride has little effect (Fig. 1c). A dependence of CIC-0 microscopic gating transitions on the Cl[−]-gradient has already been noted¹⁵.

The dependence of gating on $[\text{Cl}^-]_o$ could be due to a simple mechanism in which channel opening depends on chloride binding to a site within the pore; p_{open} should then increase with $[\text{Cl}^-]_o$ and with positive intracellular potentials. Intracellular chloride has little effect on gating, because it cannot reach its binding site in the closed state.

The simplest model assumes that chloride-binding to a single site in the pore is required for channel opening¹⁰. Although it may serve as a first approximation to our data, the shift of p_{open} with $[\text{Cl}^-]_o$ is ~20% less than minimally predicted¹⁰. This suggests that this model is either principally unsuited to explain CIC-0 gating, or that it needs further refinement.

The pores of many channels, including certain chloride channels¹⁶, can accommodate more than one ion, leading to concentration-dependent interactions within the pore. In a single-ion pore, conductance changes monotonously when the concentration ratio between two permeant ionic species is varied; in a multi-ion pore, however, interactions between different species can lead to a current minimum at a certain concentration ratio, an 'anomalous mole fraction behaviour'^{16,17}. We indeed found this effect with mixtures of Cl[−] and NO₃[−] (Fig. 2b), indicating that CIC-0 has a multi-ion pore. Thus, a model having a single chloride-binding site¹⁰ may be too simplistic.

* Deceased.

† To whom correspondence should be addressed.

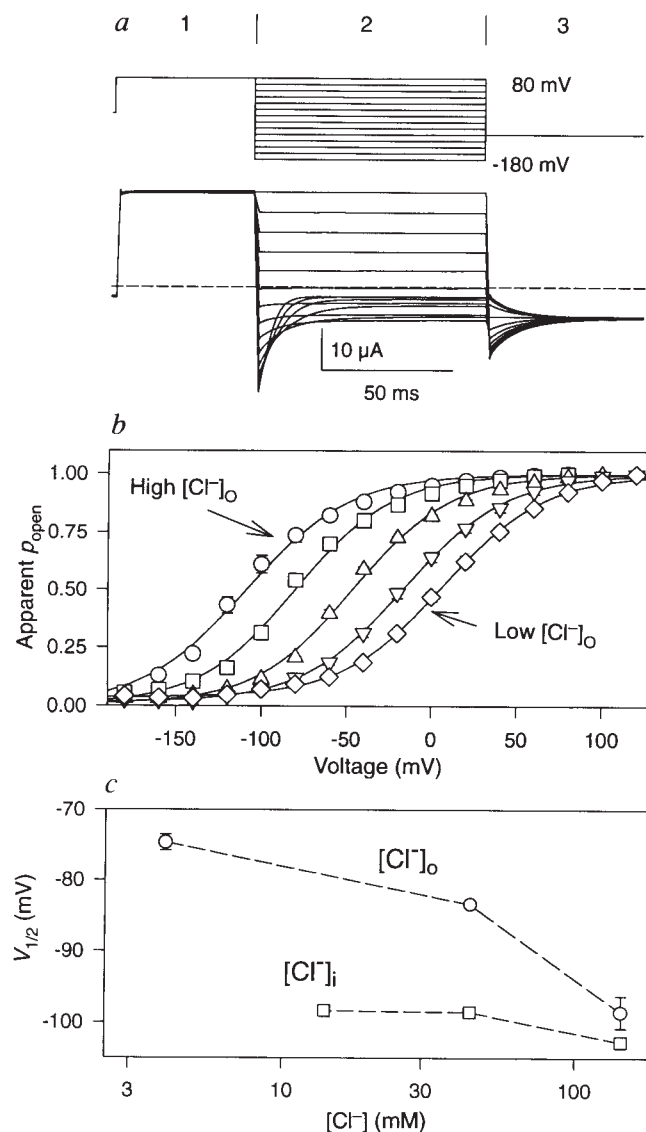
FIG. 1 Voltage- and Cl^- -dependent gating of CIC-0 (WT). **a**, Voltage-clamp protocol and current traces used to study the fast gate and the instantaneous I - V of CIC-0. To study the fast gate in isolation, the slow gate is maximally activated by previous hyperpolarization (not shown) and remains open during the pulse programme. Clamping to +80 mV (part 1) maximally opens the fast gate. Then deactivating currents are measured at different test voltages (period 2). Extrapolation to the beginning of period 2 yields instantaneous I - V of the open channel (see Fig. 4). Then p_{open} is monitored at the constant test pulse to -100 mV by extrapolating currents to the beginning of period 3. **b**, Steady-state voltage-dependence of the fast gate of CIC-0 expressed in oocytes at different $[\text{Cl}^-]_o$ (○, 100 mM; □, 21.5 mM; △, 4.6 mM; ▽, 1 mM; ◇, 0.22 mM). The high Cl^- solution contained 100 mM NaCl, 5 mM MgSO_4 and 5 mM Ca-HEPES, pH 7.4. $[\text{Cl}^-]$ was lowered by replacing it with an equimolar amount of sucrose. Gating is not affected if extracellular sodium is replaced by potassium or *N*-methyl-D-glucamine. **c**, Half-maximal activation voltage ($V_{1/2}$) of CIC-0 expressed in CHO cells at different external and internal $[\text{Cl}^-]$ (obtained by equilibration with the patch pipette solution).

METHODS. Capped cRNA was transcribed from CIC-0 (ref. 12) and injected into *Xenopus* oocytes as described^{12,4}. Currents were measured after 1–3 days using two-electrode voltage-clamp and pCLAMP5.5 software. Averaged data from 5 oocytes is shown. Apparent p_{open} was obtained by driving the fast gate into steady-state at various voltages (**a**, part 2), and then measuring the peak current (where p_{open} has not yet changed) at a constant test-potential (**a**, part 3). Gating relaxations were slow enough compared to capacitive transients to extrapolate reliably a single exponential function to the onset of the test pulse (**a**, part 3). Apparent p_{open} was normalized to the maximum current flowing after positive pre-potentials. As p_{open} is measured at a constant test potential, it is independent of single-channel I - V . CHO cells were stably transfected with CIC-0 cloned into pcDNA1-neo and selected with G418. Currents in CHO cells were measured using the whole-cell and outside-out configuration of the patch-clamp technique. External solutions were 135 mM NaCl, 5 mM KCl, 1.8 mM CaCl_2 , 1 mM MgCl_2 , 5 mM Na-HEPES. Internal solutions contained 140 mM KCl, 10 mM EGTA, 5 mM K-HEPES and 2 mM MgCl_2 . Cl^- was reduced by substitution with equal amounts of cyclamate. Data represent mean values of 2 patches for variation of external, and >10 patches for internal chloride. Error bars are s.e.m. throughout.

If the permeant anion directly gates the channel, one may expect that p_{open} will also show an anomalous mole-fraction behaviour. Indeed $V_{1/2}$ (the voltage at which $p_{\text{open}} = 0.5$) has a minimum at nearly the same mole fraction (Fig. 2b).

To identify structures involved in gating, we neutralized all charged residues in putative transmembrane spans (Fig. 3a). Except for a lysine towards the end of domain D12, the mutants could either not be expressed functionally, or were indistinguishable from wild type (WT). Gating and permeation were altered in mutants K519Q, K519E and K519H, which yielded very similar currents. K519R was intermediate to WT. Thus, a positive charge and other features of the side chain are important at this position. Mutants in K520 and in residues 515 to 518 resembled WT, demonstrating the specificity of K519. We concentrated on K519E. Its gating is slower (Fig. 3b) but still depends on $[\text{Cl}^-]_o$ (Fig. 3c). The nominal gating charge remains ~1. In contrast to WT, K519E currents do not deactivate completely at negative voltages, but the apparent p_{open} deduced from macroscopic currents saturates at ~35% (Fig. 3c). This does not result from a nonspecific leak current, but is due to a low conductance state of the channel (~0.3 pS) remaining with closed fast gates (our manuscript in preparation).

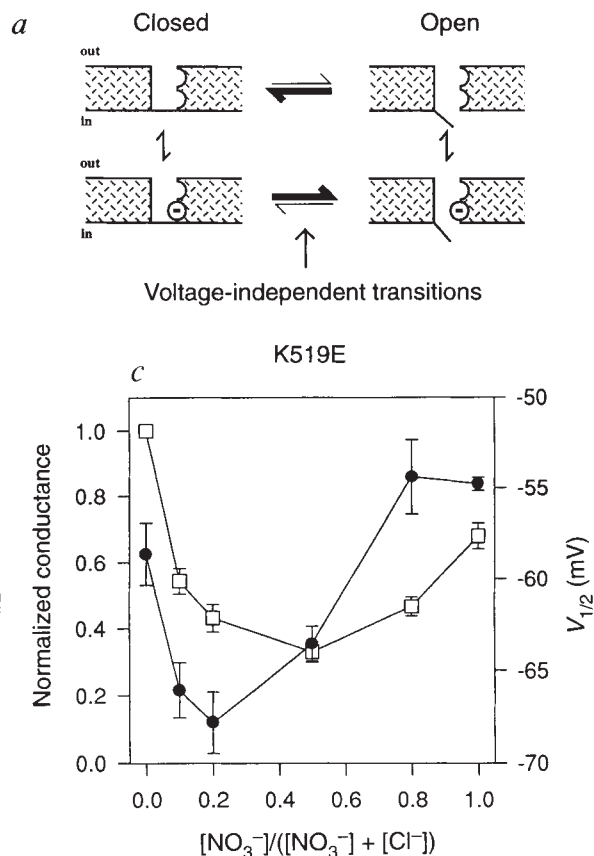
Several pore properties are also changed. Although the instantaneous I - V of WT CIC-0 Cl^- -currents is linear^{13,14}, it is outwardly rectifying with K519E (Fig. 4a, b). This agrees with the



assumed cytoplasmic location of K519 and with effects described for mutations close to pore entrances in other channels¹⁸. Noise analysis^{19,20} indicates that single-channel conductance was reduced from ~8 pS (WT) to between ~0.8 (at -140 mV) and ~3 pS (at +60 mV) for K519E, making single-channel studies unfeasible. Cl^- permeates WT CIC-0 better than Br^- and NO_3^- , and I^- blocks it (Fig. 4a). K519E, however, conducts Br^- as well as Cl^- , NO_3^- conductance is increased, and I^- block reduced (Fig. 4b). If our model is correct, one may expect that this difference in halide selectivity of conductance entails a difference in the halide selectivity of gating, because the anion must penetrate the ion-selective pore to reach its binding site. Both in WT and K519E CIC-0, gating is stimulated by Cl^- , Br^- , and NO_3^- or I^- (in comparison to glutamate or other impermeant anions) (Fig. 4c, d). This suggests that only permeant ions directly affect gating. Although WT CIC-0 is gated less efficiently (in comparison to Cl^-) with Br^- and especially I^- (Fig. 4c), this difference is indeed reduced in the less selective mutant K519E (Fig. 4d). K519E does not abolish the multi-ion behaviour, but shifts the conductance maximum to a different mole fraction (Fig. 2c). With both WT and K519E, p_{open} depends less steeply on voltage with NO_3^- as compared to Cl^- .

In voltage-dependent cation channels, pores and voltage-sensors are assigned to different structures^{8,10,17,21,23}, even though several point mutations change both gating and

FIG. 2 Schematic model of fast Cl^- -dependent gating of CIC-0 (a) and anomalous mole fraction behaviour resulting from interactions in the pore (b, c). The model in a assumes intrinsically voltage-independent gating transitions. The pore contains two anion-binding sites, and the gate is close to its cytoplasmic mouth. Occupation of the internal binding site shifts the equilibrium between closed and open conformations to the open state, irrespective of whether the outer binding site is occupied (for simplicity, occupation of outer binding site is not shown). This leads to a voltage- and concentration-dependence of gating. Binding of an anion to one site can influence the binding of a second anion to the other site (for example, by electrostatic interaction). This can account for the anomalous mole fraction effect when both NO_3^- and Cl^- are present. This interaction is also able to explain the shallower voltage-dependence of gating with NO_3^- ; in this case, the transition of NO_3^- from the outer to the inner binding site makes the major contribution to the voltage-dependence. b, c, Anomalous mole fraction behaviour of conductance ($\square$) and of gating ($\bullet$; $V_{1/2}$ is the voltage where $p_{\text{open}} = 0.5$) for WT CIC-0 (b) and the K519E mutant (c). Similar effects were found with $\text{Br}^-/\text{NO}_3^-$ mixtures, but no anomalous mole fraction behaviour was found in Cl^-/Br^- , Cl^-/I^- , or Br^-/I^- mixtures. METHODS. Channels were expressed and measured in oocytes as described in Fig. 1. The different mole fractions were obtained by replacing NaCl with equal amounts of NaNO_3 in a solution containing 100 mM NaCl and 5 mM Ca-HEPES pH 7.4. Conductance was measured as slope conductance at +60 mV. $V_{1/2}$ was obtained by Boltzmann fits to steady-



tate p_{open} . Values represent mean values from 4 oocytes. Error bars indicate s.e.m.

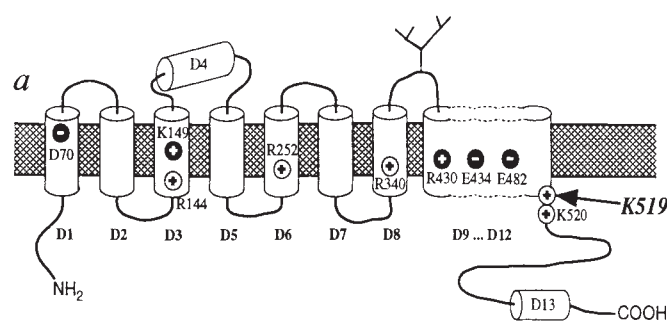
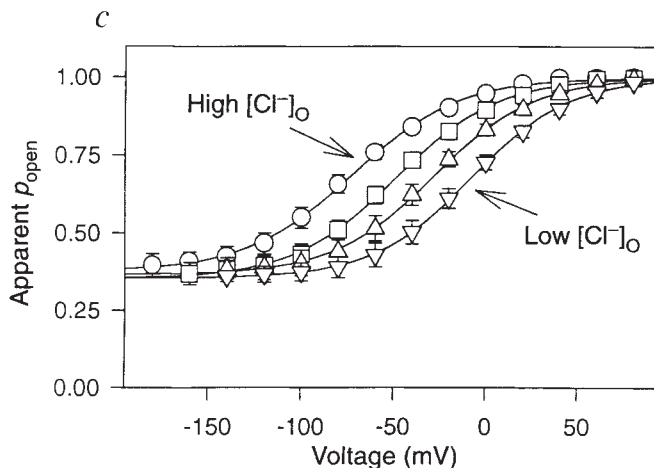
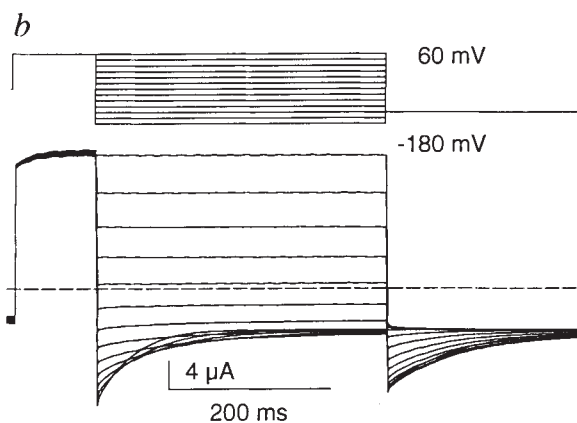


FIG. 3 Location of mutated residues (a) and properties of mutant K519E (b-c). a, Location of charged residues that were mutated to glutamine (K519 also to E, H and R). Mutants that could not be expressed functionally are indicated by filled symbols (D70, K149, R430, E434, E482). R144Q, R252Q, R340Q and K520Q were indistinguishable from WT (open symbols). K519 is conserved in CIC-0, -1, -2 and the CIC-K proteins. We also mutated non-charged residues N-terminal to K519, which gave essentially WT currents (I515L, I516L, Q517R, I518L; not shown in model). Glycosylation studies⁷, low homology in D4, and the apparent importance of K149 led us to revise the originally proposed transmembrane model^{12,5}. There may be several transmembrane spans in the broad hydrophobic region D9-D12, the exact topology of which is unknown. The region after D12 is cytoplasmic⁵. b, Voltage-clamp traces of K519E expressed in oocytes. Note the difference in time scale compared to Fig. 1a (for explanation of pulse protocol, see there). c, Steady-state p_{open} as a function of voltage and $[\text{Cl}^-]_o$. Chloride concentrations are as in Fig. 1b, except that no measurements were done with 0.22 mM Cl^- (because of the low conductance of K519E). The scale gives 'apparent p_{open} ' as calculated from macroscopic current measurements (see Fig. 1a), which differs from true p_{open} because of a residual conductance remaining with closed fast gates ('leaky closed state'). Data are averaged from 5 oocytes. METHODS. Mutations were introduced by recombinant PCR and verified by sequencing. To boost expression, K519E was cloned into a vector containing *Xenopus* globin untranslated sequences²⁹. cRNA was prepared and functionally analysed as in Fig. 1.



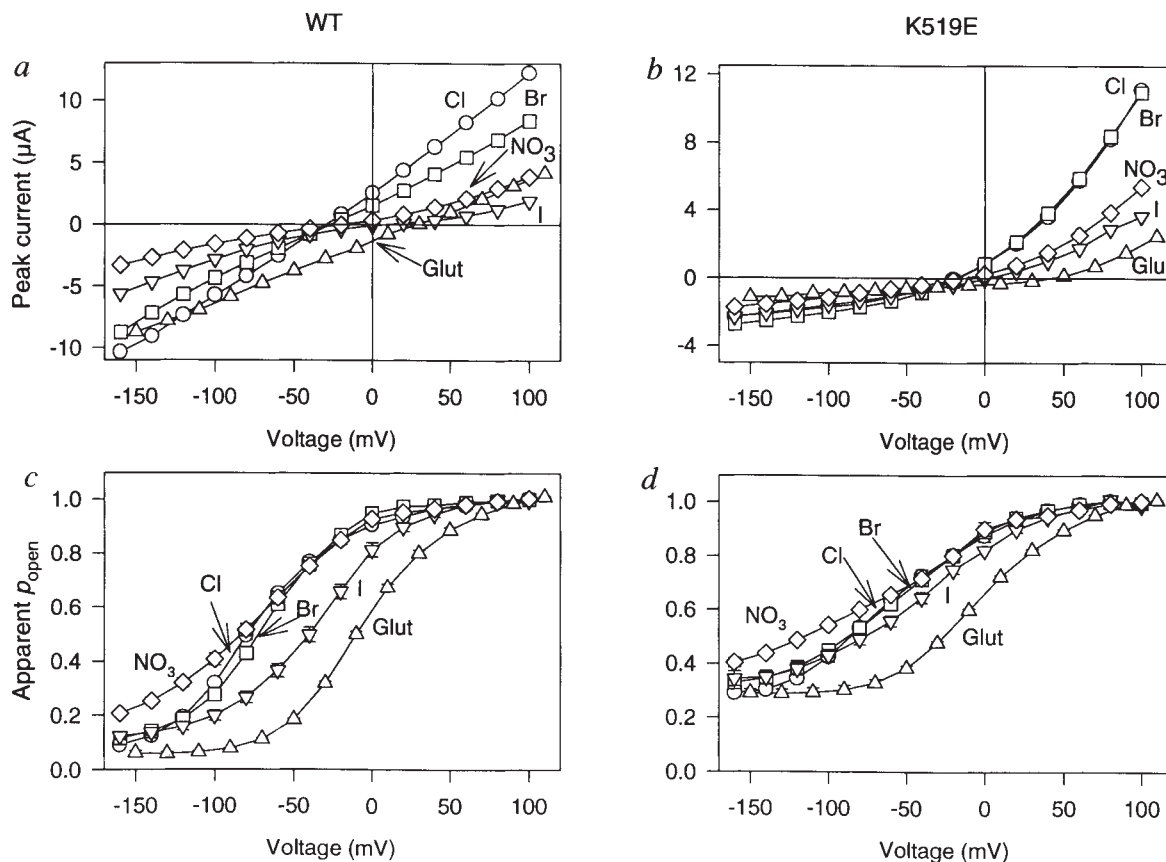


FIG. 4 Halide selectivity of conduction (a, b) and of gating (c, d) of WT and K519E CIC-0, respectively. a, b, Instantaneous (open channel) I-V of WT CIC-0 and mutant K519E, respectively, measured in ND96 (104 mM Cl⁻) (ref. 4) (○), or in ND96 in which 96 mM of Cl⁻ were substituted by Br⁻ (□), NO₃⁻ (◇), I⁻ (▽) or glutamate (△) (this leaves 8 mM Cl⁻ in the solution). Note the strong outward rectification and relative loss of selectivity of K519E. c, d, Halide selectivity of gating for WT and K519E CIC-0, respectively. Apparent p_{open} was determined as described in Fig. 1a using the solutions described for a, b. Aspartate, cyclamate, gluconate, methanesulphonate and sulphate were also unable to permeate and to increase p_{open} (results not shown). As in Fig. 3c, 'apparent

p_{open} ' of K519E saturates at ~0.35 and is different from 'true' p_{open} because of the 'leaky closed state'.

METHODS. Channels were expressed and analysed in oocytes essentially as described in Fig. 1. The linear I-V of open CIC-0 has been described^{13,14}. The outward rectification of K519E is also supported by our noise analysis (~0.8 pS at -140 mV, and ~3 pS at +60 mV, our manuscript in preparation), and by macroscopic currents in cell-attached patches which provide higher time-resolution. For a and b, representative measurements done on a single oocyte are shown; in c and d, data from 3 oocytes were averaged.

conductance^{24,25}. The influences permeant ions can have on gating of these channels are generally classified as modulatory effects on intrinsically voltage-dependent conformational changes¹⁷. This work suggests a fundamentally different gating mechanism for CIC-0, in which gating is intrinsically linked to permeation. Although a very simple model¹⁰ involving a single anion-binding site in the pore describes gating to a first approximation, the anomalous mole fraction behaviour suggests more than one binding site in the pore (Fig. 2a). Only occupation of the internal site is postulated to affect gating. Binding of chloride to either site influences binding of a second ion to the other site, explaining both the anomalous mole fraction behaviour and the shallower slope seen with NO₃⁻. Preliminary model calculations show that it can indeed describe CIC-0 gating (our manuscript in preparation).

In addition to the strong dependence of gating on [Cl⁻]_o, several findings provide independent strong support for this model. (1) Only permeant anions affect gating. (2) A specific mutation (K519E) affects both conduction and gating. (3) The ion selectivity of conduction is reflected in the ion selectivity of gating; in the mutant, ion selectivities of conduction and gating vary roughly in parallel, as expected when the gating anion has to penetrate the ion-selective pore to reach its binding site. (4) With NO₃⁻, p_{open} depends less steeply on V; this can be explained by interactions between the two binding sites. (5) With

Cl⁻/NO₃⁻ mixtures, conduction shows an anomalous mole fraction behaviour (strictly a pore property), which is reflected in a similar behaviour of gating.

Findings (3) to (5) are nearly impossible to reconcile with the most simple alternative model in which anion binding to an external site modulates intrinsically voltage-dependent gating. However, they are naturally explained in the framework of our model (Fig. 2a), for which they provide very strong support.

K519 is probably located at the intracellular end of the pore where both conduction and gating are controlled. Though tempting, it is not possible to conclude that it directly is the internal anion binding site. In agreement with our results, an alignment²⁶ of CIC-0 with vastly different cation and anion channels suggests that the D9-D12 segment may participate in permeation, though homology is very weak. Gating models similar to ours may also apply for other ion channel classes. For instance, gating of inwardly rectifying K⁺ channels, which also lack the voltage-sensor motif²⁷, depends on extracellular potassium^{27,28}. □

Received 12 August; accepted 16 December 1994.

- Steinmeyer, K. et al. *Nature* **354**, 304-308 (1991).
- Koch, M. C. et al. *Science* **257**, 797-800 (1992).
- Steinmeyer, K., Lorenz, C., Pusch, M., Koch, M. C. & Jentsch, T. J. *EMBO J.* **13**, 737-743 (1994).

4. Thiemann, A., Gründer, S., Pusch, M. & Jentsch, T. J. *Nature* **356**, 57–60 (1992).
5. Gründer, S., Thiemann, A., Pusch, M. & Jentsch, T. J. *Nature* **360**, 759–762 (1992).
6. Uchida, S. et al. *J. Biol. Chem.* **268**, 3821–3824 (1993).
7. Kieferle, S., Fong, P., Bens, M., Vandewalle, A. & Jentsch, T. J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6943–6947 (1994).
8. Stühmer, W. et al. *Nature* **339**, 597–603 (1989).
9. Papazian, D. M., Timpe, L. C., Jan, Y. N. & Jan, L. Y. *Nature* **349**, 305–310 (1991).
10. Sigworth, F. J. Q. *Rev. Biophys.* **27**, 1–40 (1994).
11. White, M. M. & Miller, C. J. *J. Biol. Chem.* **254**, 10161–10166 (1979).
12. Jentsch, T. J., Steinmeyer, K. & Schwarz, G. *Nature* **348**, 510–514 (1990).
13. Hanke, W. & Miller, C. J. *gen. Physiol.* **82**, 25–45 (1983).
14. Bauer, C. K., Steinmeyer, K., Schwarz, J. R. & Jentsch, T. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11052–11056 (1991).
15. Richard, E. A. & Miller, C. *Science* **247**, 1208–1210 (1990).
16. Tabcharani, J. A. et al. *Nature* **366**, 79–82 (1993).
17. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 1992).
18. Imoto, K. et al. *Nature* **335**, 645–648 (1988).
19. Heinemann, S. H. & Conti, F. *Meth. Enzym.* **207**, 131–148 (1992).
20. Pusch, M., Steinmeyer, K. & Jentsch, T. J. *Biophys. J.* **66**, 149–152 (1994).
21. Yool, A. J. & Schwarz, T. L. *Nature* **349**, 700–704 (1991).
22. Hartmann, H. A. et al. *Science* **251**, 942–944 (1991).
23. Heinemann, S. H., Terlau, H., Stühmer, W., Imoto, K. & Numa, S. *Nature* **356**, 441–443 (1992).
24. Kirsch, G. E. et al. *Neuron* **8**, 499–505 (1992).
25. Heglinbotham, L., Lu, Z., Abramson, T. & MacKinnon, R. *Biophys. J.* **66**, 1061–1067 (1994).
26. Jan, L. Y. & Jan, Y. N. *Nature* **371**, 119–122 (1994).
27. Kubo, Y., Baldwin, T. J., Jan, Y. N. & Jan, L. Y. *Nature* **362**, 127–133 (1993).
28. Hagiwara, S., Miyazaki, S. & Rosenthal, N. P. *J. gen. Physiol.* **67**, 621–638 (1976).
29. Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057–7070 (1984).

ACKNOWLEDGEMENTS. We thank C. Schmekal for technical assistance. This work was supported, in part, by the Bundesministerium für Forschung und Technologie, the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie and the US Cystic Fibrosis Foundation.

Expression of *relB* is required for the development of thymic medulla and dendritic cells

Linda Burkly*, Catherine Hession*, Lynn Ogata†, Christina Reilly†, Lori Anne Marcon†, Dian Olson*, Richard Tizard*, Richard Cate* & David Lo††

* Biogen Inc., Cambridge, Massachusetts 02142, USA

† Department of Immunology, The Scripps Research Institute, La Jolla, California 92037, USA

DENDRITIC cells (DC) derived from bone marrow are critical in the function of the immune system, for they are the primary antigen-presenting cells in the activation of T-lymphocyte response. Their differentiation from precursor cells has not been defined at a molecular level, but recent studies have shown an association between expression of the *relB* subunit of the NF- κ B complex^{1–5} and the presence of DC in specific regions of normal unstimulated lymphoid tissues^{4–6}. Here we show that *relB* expression also correlates with differentiation of DC in autoimmune infiltrates *in situ*, and that a mutation disrupting the *relB* gene results in mice with impaired antigen-presenting cell function, and a syndrome of excess production of granulocytes and macrophages. Thymic *UEA-1*⁺ medullary epithelial cells from normal mice show striking similarities to DC and, interestingly, these cells are also absent in *relB* mutant mice. Taken together, these results suggest that *relB* is critical in the coordinated activation of genes necessary for the differentiation of two unrelated but phenotypically similar cells (DC and thymic *UEA-1*⁺ medullary epithelial cells) and is therefore a candidate for a gene determining lineage commitment in the immune system.

Although *relB* expression can be found in B-cell enriched fractions^{4,5}, recent studies using *in situ* hybridization and immunostaining showed stronger expression specifically associated with DC in lymphoid tissues⁶, suggesting a role for *relB* in the development of DC from bone marrow precursors. Consist-

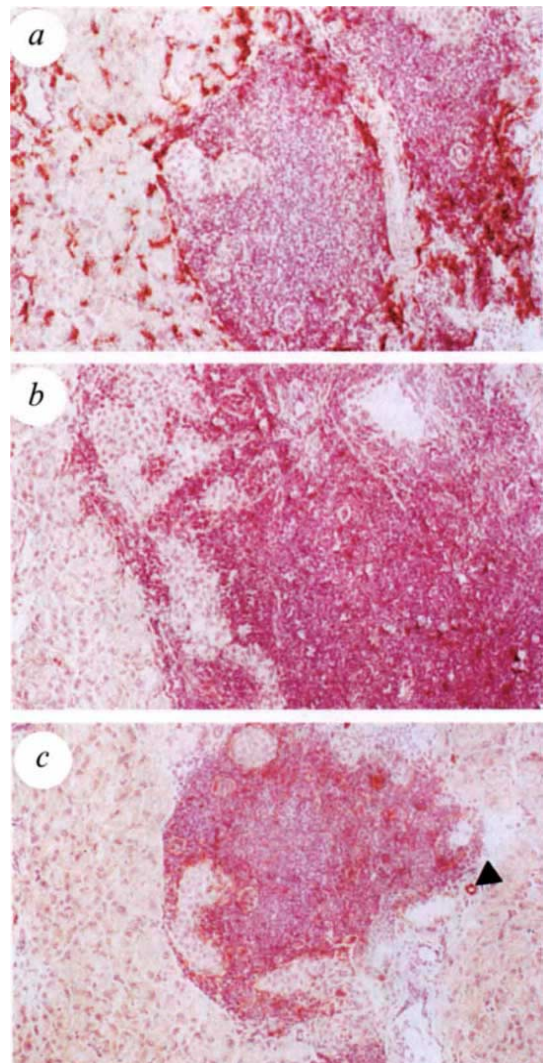


FIG. 1 Histology of inflammation in pancreatic islets. *a*, Stained for F4/80 (macrophages), showing their distribution at the periphery of the islet infiltrate; *b*, stained for *relB*, showing colocalization with dendritic cells in the islet infiltrate; *c*, for VCAM-1 showing network of dendritic cells in islet infiltrate, and VCAM-1⁺ activated vascular endothelium (arrowhead).

METHODS. Diabetic mice carrying both the TCR-HNT and Ins-HA transgenes have been described previously. These mice express an islet β -cell-specific transgene antigen (influenza haemagglutinin, or HA) and T-cell receptors specific for HA, and develop early onset spontaneous autoimmune diabetes²⁷ with infiltrates organized by VCAM-1⁺ DC²⁸. Immunostaining of cryostat sections have been described previously²⁷; briefly, cryostat sections of pancreas from the transgenic mice were fixed briefly in 1% paraformaldehyde, incubated with primary antibodies (rat monoclonals F4/80, or anti-VCAM-1 (M/K-2), and rabbit polyclonal anti-*RelB* (Santa Cruz Biotechnology)), followed by secondary antibodies against the primary (biotin-anti-rat Ig for F4/80 and VCAM-1, followed by avidin peroxidase; peroxidase-conjugated anti-rabbit Ig for the *RelB*), incubation with amino-ethyl-carbazole (AEC) and haematoxylin counterstain. Magnification, $\times 160$.

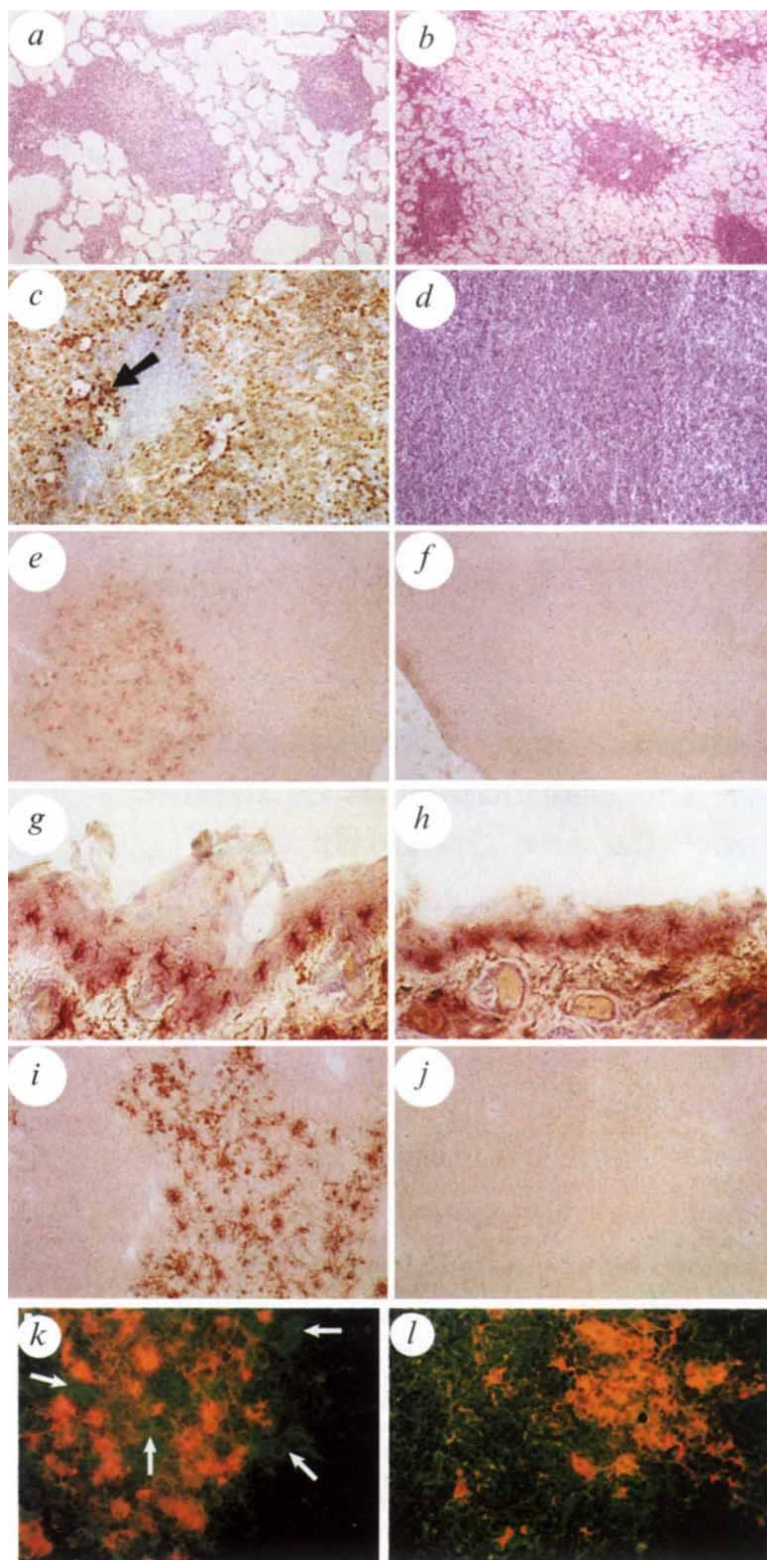
ent with this, analysis of islet lymphoid infiltrates in a model of autoimmune diabetes revealed expression of *relB* within the DC-containing regions (Fig. 1*b, c*) and not among the functionally related macrophages (Fig. 1*a*), suggesting that *relB* expression correlates with new induction of DC differentiation.

Direct evidence for a relationship between *relB* and DC differentiation was found in mice with an insertional mutation in the *relB* gene, inactivating its expression. This mutation was caused by chance integration of a class I major histocompat-

† To whom correspondence should be addressed.

FIG. 2 Histology of mutant, and expression of *relB* in hemizygotes, mutants and chimaeras. **a–d**, Histology of 272H mutant tissues. **a**, Periodic acid–schiff (PAS) staining showing granulocyte and monocyte infiltrates in lung parenchyma, sparing alveolar spaces. **b**, 272H liver, showing infiltrates around portal vessels, and extensive fibrosis throughout parenchyma (PAS). **c**, 272H spleen, stained for F4/80⁺ macrophages, showing sparsely populated white pulp, large numbers of F4/80⁺ macrophages in the red pulp, and excessive granulocytopenia (granular staining by endogenous peroxidase, arrow). **d**, 272H thymus, showing absence of medulla-like regions (haematoxylin). **e**, 272 hemizygote thymus stained for RelB, showing expression in medulla, and weak staining in occasional cortical cells. **f**, 272H mutant (littermate to **e**) thymus, stained for RelB, showing absence of staining. **g**, Control ear skin, stained for class II MHC, showing staining of Langerhans cells in the dermis–epidermis boundary. **h**, 272H mutant ear skin, stained for class II MHC, showing normal development of Langerhans cells. **i–l**, Thymus medulla development in normal, mutant and 272H→B6 chimaera. **i**, Hemizygote thymus stained for medullary thymic epithelium marker UEA-1. **j**, 272H mutant thymus stained for UEA-1. **k**, Normal thymus double stained for RelB (green) and UEA-1 (red). Bone marrow-derived thymic dendritic cells appear as single staining (green) cells (arrows), whereas UEA-1⁺ MEC appear as double-staining (yellow) cells. **l**, 272H→B6 chimaera thymus stained as for **k**, showing only the host radiation-resistant UEA-1⁺ MEC as double-staining (yellow) cells.

METHODS. Immunostaining was done as described in Fig. 1, but UEA-1 staining was done using biotin-UEA-1, followed by avidin-peroxidase. **e, f, i, j**, Shown without counterstain so that specific staining would not be obscured. For **k, l**, the enzymatic detection steps were replaced by FITC-conjugated anti-Rabbit Ig (for RelB), and avidin-Rhodamine (for UEA-1). 272H→B6 chimaeras were made by injecting 5×10^6 bone marrow cells from 272H donors into lethally irradiated (1100R) C57BL/6 recipients (the original 272 line was generated on the inbred C57BL/6 strain). Flow cytometry confirmed that, under these conditions, essentially all of the lymphoid cells expressed the donor class I K^k transgene. Magnifications: **a, b, d**, $\times 80$; **c, e, f, i, j**, $\times 160$; **g, h**, $\times 320$; and **k, l**, $\times 225$.



ibility complex (MHC) genomic clone as a transgene ('line 272', ref. 7); mice homozygous for the 272 transgene integrant ('272H') showed a severe inflammatory syndrome, with mixed granulocyte/monocyte infiltrates in lung, liver and other tissues, immune system defects, and absence of the thymic medulla (Fig. 2a–d). Cloning of the region flanking the transgene integrant revealed disruption of the *relB* gene just downstream from the exon encoding the *rel* domain (Fig. 3a–d), separating it from C-terminal sequences required for transactivator activity³. Mapping of the integrant accounted for all transgene sequences as a

single contiguous integrant, ruling out multiple integration events. Southern blot analysis (Fig. 3c) also demonstrated that downstream *relB* sequences remained intact, but were contained in a novel restriction fragment in the genome of 272 hemi- and homozygous mutant mice.

The disrupted *relB* gene was found to constitute a recessive loss-of-function mutation by two criteria. First, reverse transcribed-polymerase chain reaction (RT-PCR) analysis of spleen and thymus for *relB* expression showed that messenger RNA transcripts encoding the *relB* gene were not found in 272H

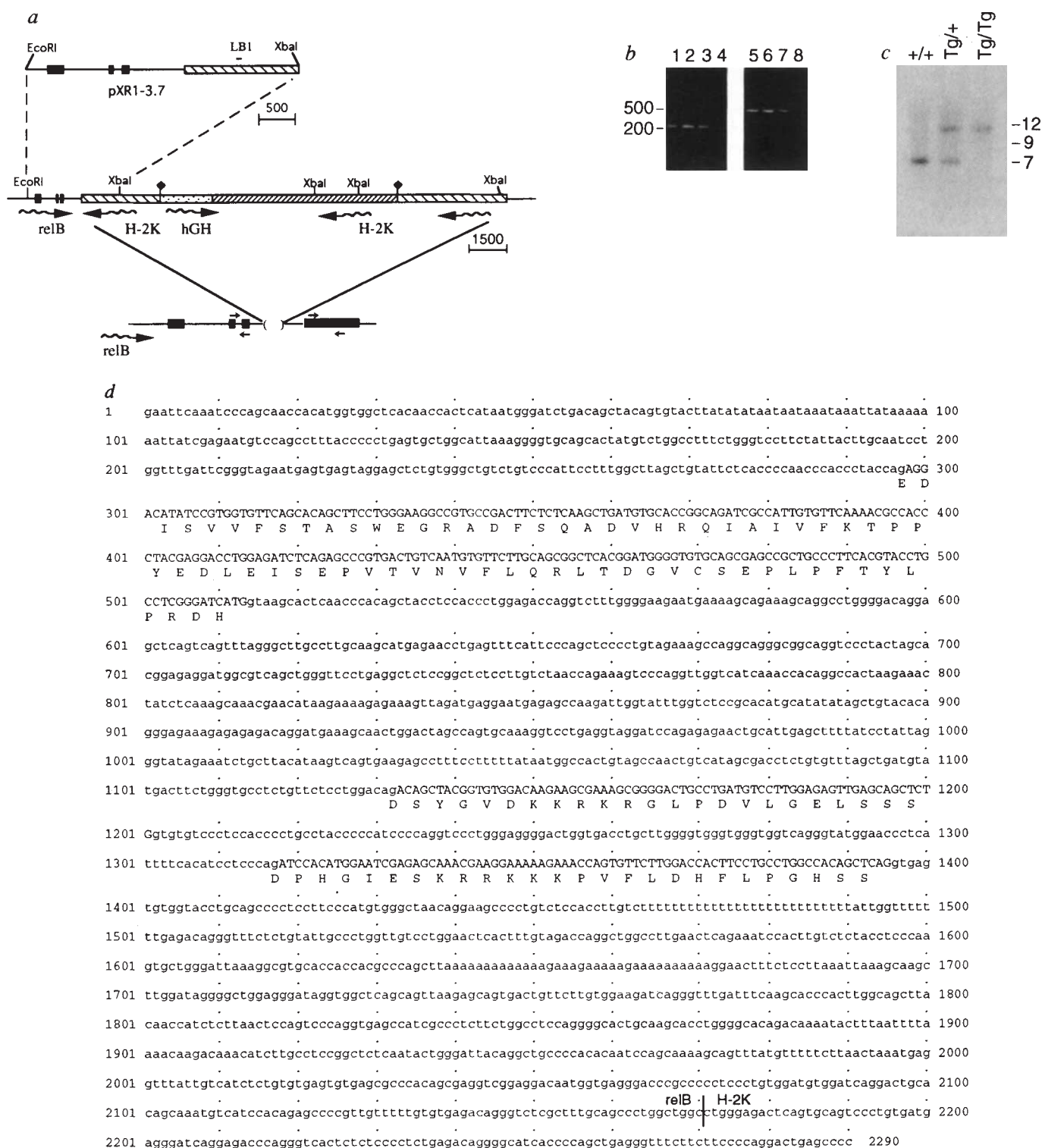


FIG. 3 Characterization of the mutant *reIB* locus and expression analysis. **a**, Disruption of the *reIB* locus by transgene sequences (18 kb) is depicted, including two copies of genomic H-2K^k (one intact copy and one broken copy shown as hatched boxes, the 5' end of each copy indicated by a diamond), and one copy of an hGH gene fragment (stippled box). Transcriptional direction of genes is indicated by wavy arrows. A 3.7 kb *EcoRI*-*XbaI* fragment, pXR1-3.7, was cloned using a transgene-specific oligonucleotide probe, LB1. The transgene inserted within a *reIB* intron of undefined length (parentheses). 3' *reIB* exon sequences identified in homozygous mutant DNA are schematically shown as a single exon based on PCR analysis. **b**, PCR analysis showing intact *reIB* 5' (lanes 1-3) and 3' (lanes 5-7) sequences in normal C57BL/6 (lanes 1, 5), 272 hemizygote (lanes 2, 6), and 272 Homozygote (lanes 3, 7) genomic DNA. The 240 bp band for 5' sequences

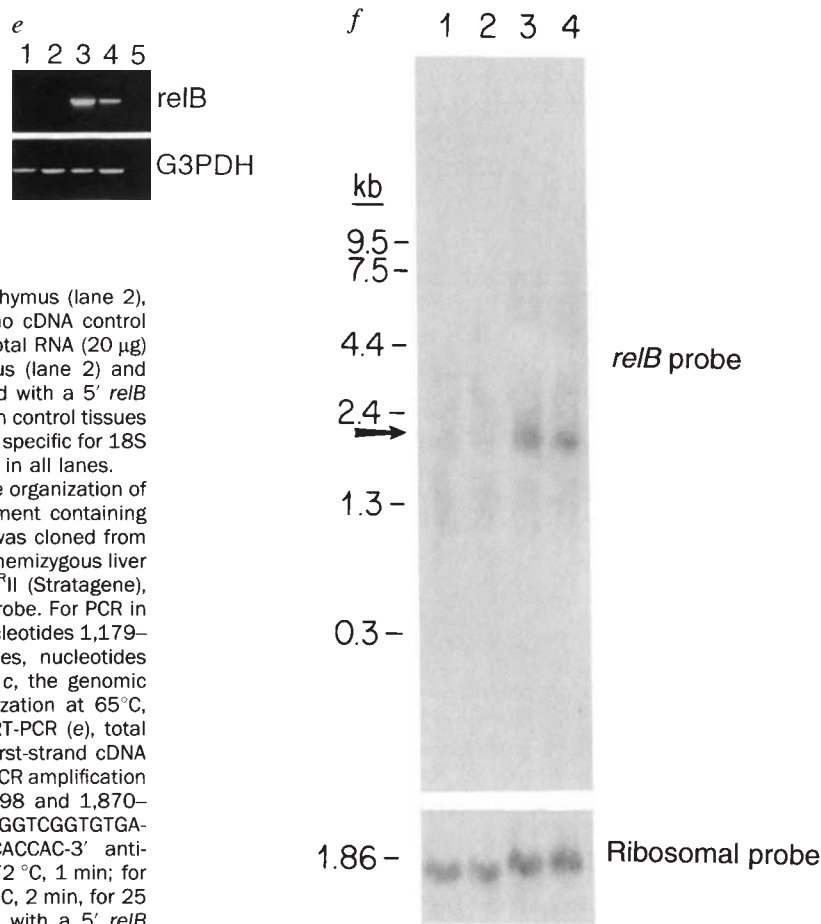
includes 115 bp intron. Lanes 4 and 8, no template control. **c**, Southern blot analysis showing a restriction-fragment length polymorphism between the wild-type and mutant *reIB* locus detected by a 3' *reIB* probe, confirming both disruption of *reIB* and retention of 3' *reIB* sequence. *Bam*HI-digested genomic DNA of C57BL/6 (+/+), 272 hemizygous (Tg/+) and homozygous (Tg/Tg) mice was hybridized with a 3' *reIB* fragment (corresponding to nucleotides 1,361-1,889 of the *reIB* cDNA, a region downstream from the transgene integrant). The +/+ and Tg/+ DNA exhibited a 7 kb fragment, whereas Tg/+ and Tg/Tg DNA exhibited a novel 10-11 kb fragment. **d**, Sequence of the pXR1-3.7 clone. *reIB* exon sequence (upper case) with corresponding amino acid sequence (below) and intron sequences (lower case) are shown. The *reIB* exons from the cloned genomic region encode the *rel* homology domain (nucleotides 959-1,174 of the published cDNA sequence³) and

(Fig. 3 legend continued from page 533)

nuclear localization sequences (1,175–1,243 and 1,244–1,321). At the 5' junction with *relB*, H-2K^b regions in pXR1-3.7 correspond to nucleotides 2,106–554 of the published sequence²⁹. e, RT-PCR analysis of *relB* expression in mutant and normal tissues. Primers for *relB* corresponding to sequences 5' and 3' of the transgene junctions amplified a 711 bp *relB* cDNA fragment from normal but not homozygous mutant tissues. A G3PDH cDNA control fragment was amplified in both.

Shown are homozygous mutant spleen (lane 1) and thymus (lane 2), normal spleen (lane 3), and thymus (lane 4) and a no cDNA control (lane 5). f, Northern blot analysis of *relB* expression. Total RNA (20 µg) from homozygous mutant spleen (lane 1) and thymus (lane 2) and control spleen (lane 3) and thymus (lane 4) hybridized with a 5' *relB* probe shows the 2.2 kb *relB* transcript (arrow) present in control tissues but not in mutant tissues. Rehybridization with a probe specific for 18S ribosomal RNA shows comparable 1.86 kb RNA bands in all lanes.

METHODS. Detailed Southern blot analysis revealed the organization of the transgenic DNA and identified an *EcoRI*–*XbaI* fragment containing flanking genomic DNA. A 3.7 kb fragment, pXR1-3.7, was cloned from a genomic library made with *EcoRI*–*XbaI*-digested 272 hemizygous liver DNA, size fractionated (3–6 kb), constructed in λ ZAP[®]II (Stratagene), and screened with an H-2K^b-specific oligonucleotide probe. For PCR in b (noted by arrows in a), primers defining *relB* cDNA nucleotides 1,179–1,198 and 1,280–1,303 were used for 5' sequences, nucleotides 1,361–1,380 and 1,870–1,889 for 3' sequences. In c, the genomic Southern blot was probed as described, with hybridization at 65°C, washing to a stringency of 0.5 × SSC, 1% SDS. For RT-PCR (e), total RNA was prepared from normal and mutant tissue. First-strand cDNA synthesis was done using oligo-(dT) primer followed by PCR amplification with primers corresponding to nucleotides 1,179–1,198 and 1,870–1,889 of the *relB* cDNA and G3PDH primers (5'-TGAAGTCGGTGTGA-ACGAGTTGGC-3' sense, 5'-CATGTAGCCATGAGGTCCACCAC-3' anti-sense). PCR cycles were 94°C, 1 min; 55°C, 1 min; 72°C, 1 min; for 30 cycles for *relB*, and 94°C, 1 min; 60°C, 1 min; 72°C, 2 min, for 25 cycles for G3PDH. The northern blot (f) was probed with a 5' *relB* fragment (240 bp) amplified from genomic DNA as shown in b with hybridization at 65°C and washing to 0.5X SSC, 1% SDS. The oligo



nucleotide specific for 18S ribosomal RNA was 5'-GGGCGGTGTGTACAA-AGGGCAGGA-3'.

mutant mice (Fig. 3e). Northern blot analysis using a *relB* 5' region probe also failed to detect aberrant transcripts or splice products (Fig. 3f), consistent with the fact that the mutant phenotype was recessive, and not due to any dominant negative effects of the mutation or transgene expression. Second, immunohistochemical staining with an antiserum raised against a C-terminal peptide encoded by *relB* showed no detectable protein in thymus medulla (Fig. 2e,f), spleen white pulp or any other tissue of 272H mice (not shown), whereas hemizygous littermates were normal.

DC are thought to be the principal stimulators of T-lymphocyte proliferation in mixed lymphocyte cultures⁸. To test this function in the *relB* mutant, titrations of irradiated spleen antigen-presenting cells (APC) from transgenic hemizygous mice, homozygous mutant mice, and normal non-transgenic stimulators were cultured with T-lymphocyte responders (Fig. 4a–c). Titrations of transgene hemizygous stimulators (H-2^b plus transgene H-2K^b) showed strong stimulation of allogeneic B10.D2 (H-2^d) and C57BL/6 (H-2^b) T cells, but homozygous mutant cells showed extremely poor (but not negative) stimulation. Thus, the *relB* mutation appears to have impaired APC function among spleen cells. The residual stimulation was probably due to the abundant macrophages (ref. 7, Fig. 2c), which have at least a 5–100-fold reduced potency as stimulators relative to DC^{8–10}.

Normal mice and 272 hemizygotes infected with influenza A/PR8 virus generate strong IgG2a antibody responses against the influenza haemagglutinin (HA; refs 11, 12 and Fig. 4d). Thus, assay of IgG antibody responses to HA in response to infection would be a good measure of integrated helper T-cell responses and B-cell maturation. Mutant mice were often sickly, so we generated bone marrow chimaeras using mutant bone marrow

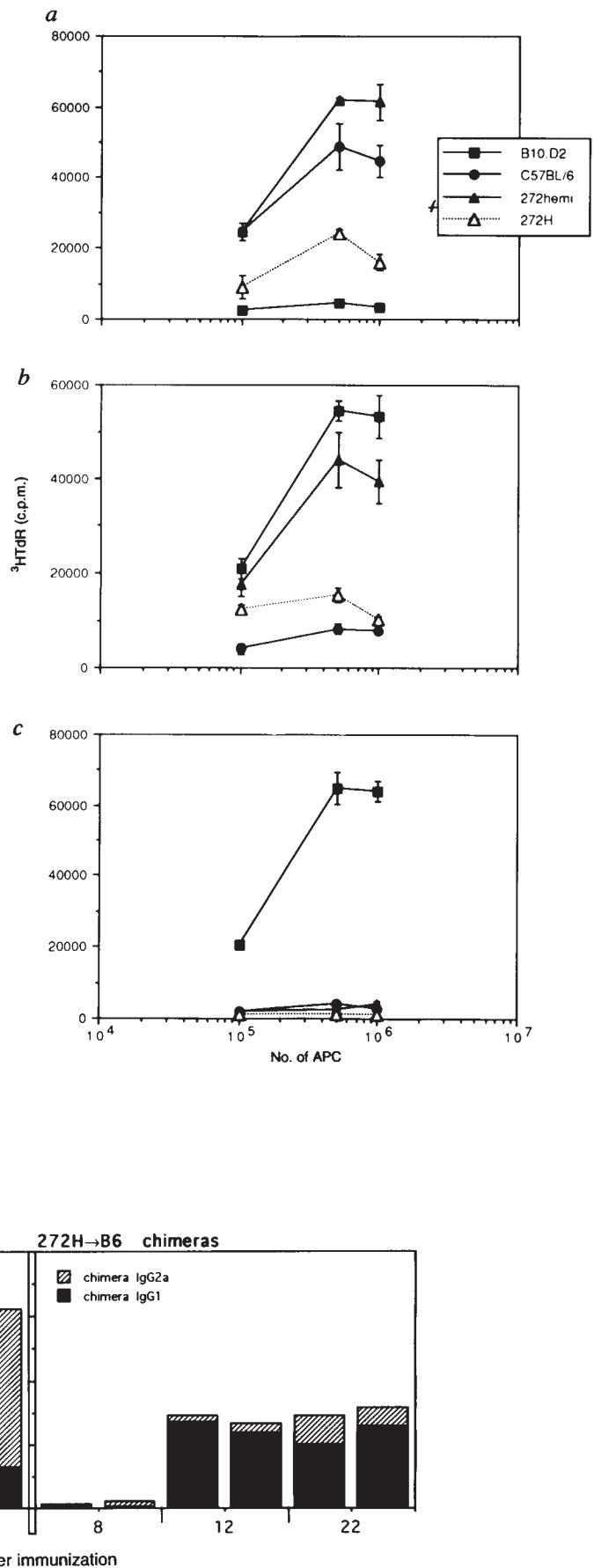
donors and irradiated normal C57BL/6 recipients ('272H→B6 chimaeras'). Such chimaeras show a less severe form of the mutant syndrome⁷, but as with the 272H mice, spleen APC from 272H→B6 chimaeras showed significantly reduced ability to stimulate allogeneic T-cell proliferation (not shown). The apparent absence of DC in 272H→B6 chimaeras did not entirely abrogate immune responses to influenza virus infection; virus-infected chimaeras generated significant (albeit lower) overall antibody titres against HA, but responses were dominated by IgG1 rather than IgG2a. This dominance of IgG1 suggests that DC are required for Th1 helper responses (producing interferon-γ) driving IgG2a antibody production, whereas macrophages or B cells as APC promote Th2 helper responses (producing interleukin-4) driving IgG1 and IgE production^{13,14}. An alternative explanation is that an absence of *relB* function in mutant B cells has subtle effects on isotype switching⁵.

Despite histological and functional evidence for the absence of DC in 272H mutant mice, it was not clear whether the deficit involved only mature DC, or the entire lineage, including precursors. Skin Langerhans cells (LC) are considered to be a possible immediate DC precursor, as strong APC function is rapidly inducible in these cells on activation by granulocyte macrophage colony-stimulating factor (GM-CSF)^{15,16}. LC in normal mouse skin were *relB* negative by immunostaining, and LC were present in normal numbers in mutant mice (class II⁺ VCAM-1⁺ F4/80⁺ *relB*[−]; Fig. 2g, h, and data not shown). This finding narrows the APC defect in 272H mutant mice to fully mature DC.

As noted above, 272H mice had little or no thymic medulla. Although this could have been consistent with a loss of only DC, there was also an absence of other cells associated with the thymic medulla, most notably epithelial cells that bind the lectin UEA-1 (UEA-1⁺ medullary epithelial cells (MEC); refs 17, 18).

FIG. 4 APC function in 272H mice and 272H→B6 chimaeras. **a**, B10.D2 ($H-2^d$) responder lymph node cells were cultured with titrations of irradiated spleen APC from syngeneic B10.D2, allogeneic C57BL/6 ($H-2^b$), 272 hemizygous (class I K^k transgene on an inbred C57BL/6 background), and 272 homozygous mice (272H). Although strong responses were evident against C57BL/6 and 272 hemizygous APC, only weakly positive stimulation was found against the 272H APC. **b**, C57BL/6 responders were stimulated as in **a**, with expected responses against transgene class I K^k on 272 hemizygous APC and allogeneic B10.D2, but only weak responses were generated against 272H APC stimulators. **c**, 272 hemizygous responders stimulated as in **a**, showing expected response against B10.D2 APC, but no response to syngeneic C57BL/6, 272 hemizygous APC and 272H APC. **d**, Antibody responses to influenza haemagglutinin from 272 hemizygotes and 272H→B6 chimaeras injected intraperitoneally with A/PR8 virus on days 0 and 8. Mice were bled on days 0, 8, 12 and 22, and serum was tested for IgG1 and IgG2a titres against HA. The data represent titres of anti-HA IgG of the specified isotype from two individual 272 hemizygotes and two chimaeras, and is shown as individual measurements from each animal. Although both types of animals responded to infection, normal 272 hemizygotes produced the expected IgG2a response, but the chimaeras produced mostly IgG1. Anti-HA titres were undetectable from preimmune bleeds, and so are not shown here.

METHODS. In **a–c**, mixed lymphocyte cultures were as follows: 2×10^5 responder lymph node cells were cultured with titrated numbers of irradiated (2100R) spleen stimulators for 4 days. During the last 24 h of culture, $^3\text{HTdR}$ was added, and incorporation of label is reported in the figure as c.p.m. In **d**, serum from immunized mice was diluted, and incubated with *Drosophila* cells transfected with a plasmid expressing the influenza haemagglutinin driven by the metallothionein promoter (L.O. *et al.*, unpublished results). Binding of mouse antibody was detected by Ig isotype-specific secondary antibodies conjugated to biotin, followed by avidin-conjugated alkaline phosphatase, and bound enzymatic activity was assayed by fluorescence of the reaction product from 4-methylumbelliferylphosphate. Titration curves were calibrated by comparisons between uninduced fly cells and CuSO_4 -induced cells, bound by known monoclonal IgG1 and IgG2a anti-HA antibodies.



These epithelial cells (keratin positive, radiation resistant) show a number of remarkable similarities to DC derived from bone marrow: they express high levels of class I and class II MHC antigens, the CD28 ligand B7 (refs 19, 20) and can present antigen to T-cell clones *in vitro*²¹. *UEA-I*⁺ MEC have also been shown to be capable of mediating negative selection of developing thymocytes^{19,22,23}, but not positive selection²⁴. So although *UEA-I*⁺ MEC are not bone marrow derived, it was striking that these cells were also absent in *relB* mutant mice (Fig. 2i,j). Double staining of thymus confirmed the expression of *relB* on both thymic dendritic cells and *UEA-I*⁺ MEC; although normal thymus contained both *relB*⁺ *UEA-I*⁺ dendritic cells and *relB*⁺ *UEA-I*⁺ MEC (Fig. 2k), 272H→B6 chimaeric thymuses contained only radiation-resistant *relB*⁺ *UEA-I*⁺ MEC from the normal host thymus (Fig. 2l). Overall, these results suggest that *relB* is also necessary for the development of *UEA-I*⁺ MEC, and it is the development of these cells (and not DC) that determine the development of the thymic medulla.

As shown here, the expression *relB* correlates with DC differentiation *in situ*, and disruption of *relB* expression blocked the development of DC and thymic *UEA-I*⁺ MEC. That *relB* is required for the development of similar phenotypes in cells with disparate origins strongly suggests that *relB* directly regulates expression of genes associated with the differentiated phenotype, although demonstration that *relB* is by itself sufficient to trigger DC or *UEA-I*⁺ MEC development will require transfection of precursor cells with *relB*. Further studies on the regulation of DC development may also explain the syndrome observed in mutant mice. For example, macrophages, granulocytes and DC are thought to derive from a common precursor²⁵; homeostatic mechanisms may drive these precursors to compensate for the DC deficit in mutant mice, and promote production of macrophages and granulocytes. Thus, the present *relB* mutant now provides an important new system for detailed molecular studies on development and lineage commitment in the immune system, and provides confirmation of the prediction that, as with insects, Rel-containing genes are involved both in development and immunity²⁶. □

Catalytic specificity of protein-tyrosine kinases is critical for selective signalling

Zhou Songyang^{*†}, Kermit L. Carraway III^{*}, Michael J. Eck[‡], Stephen C. Harrison[‡], Ricardo A. Feldman[§], Moosa Mohammadi^{||}, Joseph Schlessinger^{||}, Stevan R. Hubbard[¶], Darrin P. Smith[☆], Charis Eng^{☆☆}, Marla J. Lorenzo[☆], Bruce A. J. Ponder[☆], Bruce J. Mayer^{††‡} & Lewis C. Cantley^{*†‡}

^{*} Division of Signal Transduction, Department of Medicine, Beth Israel Hospital and Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02215, USA

[†] Department of Physiology, Tufts University School of Medicine, Boston, Massachusetts 02111, USA

[‡] Howard Hughes Medical Institute and Laboratory of Molecular Medicine, Children's Hospital, Boston, Massachusetts 02115, USA

[§] Department of Microbiology and Immunology, University of Maryland School of Medicine, and Medical Biotechnology Center, University of Maryland Biotechnology Institute, Baltimore, Maryland 21201, USA

^{||} Department of Pharmacology, New York University Medical Center, New York, New York 10016, USA

[¶] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA

[☆] CRC Human Cancer Genetics Research Group, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

^{☆☆} Department of Medicine, Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

^{††} Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

How do distinct protein-tyrosine kinases activate specific downstream events? Src-homology-2 (SH2) domains on tyrosine kinases or targets of tyrosine kinases recognize phosphotyrosine in a specific sequence context and thereby provide some specificity¹⁻³. The role of the catalytic site of tyrosine kinases in determining target specificity has not been fully investigated. Here we use a degenerate peptide library to show that each of nine tyrosine kinases investigated has a unique optimal peptide substrate. We find that the cytosolic tyrosine kinases preferentially phosphorylate peptides recognized by their own SH2 domains or closely related SH2 domains (group I; ref. 3), whereas receptor tyrosine kinases preferentially phosphorylate peptides recognized by subsets of group III SH2 domains³. The importance of these findings for human disease is underscored by our observation that a point mutation in the RET receptor-type tyrosine kinase, which causes multiple endocrine neoplasia type 2B, results in a shift in peptide substrate specificity.

We developed a new technique to study the substrate preference of protein kinases using an oriented peptide library⁴. In this procedure, a consensus sequence of preferred substrates is obtained by sequencing the phosphopeptide products. We used this technique to determine the optimal peptide substrates for nine protein-tyrosine kinases (PTKs) (Table 1) and found that each PTK has its own optimal substrate. Most of the PTKs selected peptides with Glu or Asp residues at specific sites N-terminal of the Tyr, in agreement with previous studies⁵⁻⁸. But selectivity was usually dominated by a preference for specific hydrophobic amino acids at key positions: the cytosolic PTKs selected substrates with Ile or Val at the -1 position and a Glu, Gly or Ala at the +1 position. The selection for Ile and Val at

Received 13 October; accepted 20 December 1994.

- Grilli, M., Chiu, J. S. & Lenardo, M. J. *Int. Rev. Cytol.* **143**, 1-62 (1993).
- Nolan, G. P. & Baltimore, D. *Curr. Opin. Genet. Dev.* **2**, 211-220 (1992).
- Ryseck, R. P. et al. *Molec. cell. Biol.* **12**, 674-684 (1992).
- Lernbecher, T., Müller, U. & Wirth, T. *Nature* **365**, 767-770 (1993).
- Lernbecher, T., Kistler, B. & Wirth, T. *EMBO J.* **13**, 4060-4069 (1994).
- Carrasco, D., Ryseck, R. P. & Bravo, R. *Development* **118**, 1221-1231 (1993).
- Lo, D., Quill, H., Burky, L., Scott, B., Palmer, R. D. & Brinster, R. L. *Am. J. Pathol.* **141**, 1237-1246 (1992).
- Metlay, J. P., Pure, E. & Steinman, R. M. *J. exp. Med.* **169**, 239-254 (1989).
- Sprent, J. & Schaefer, M. *Int. Immun.* **1**, 517-525 (1989).
- Levin, D., Constant, S., Pasqualini, T., Flavell, R. & Bottomly, K. *J. Immun.* **151**, 6742-6750 (1993).
- Coutelier, J. P., van der Logt, J. T. M., Heessen, F. W. A., Warnier, G. & van Snick, J. *J. exp. Med.* **165**, 64-69 (1987).
- Hocart, M. J., MacKenzie, J. S. & Stewart, G. A. *J. gen. Virol.* **70**, 809-818 (1989).
- Janeway, C. A. et al. *Immun. Rev.* **101**, 39-80 (1988).
- Mosmann, T. R. & Coffman, R. L. *Rev. Immun.* **7**, 145-173 (1989).
- Knight, S. C. & Stagg, A. *J. Curr. Opin. Immun.* **5**, 374-382 (1993).
- Larsen, C. P. et al. *J. exp. Med.* **172**, 1483-1493 (1990).
- Farr, A. G. & Anderson, S. K. *J. Immun.* **134**, 2971-2979 (1985).
- Surh, C. D. et al. *J. exp. Med.* **176**, 495-505 (1992).
- Degermann, S., Surh, C. D., Glimcher, L. H., Sprent, J. & Lo, D. *J. Immun.* **152**, 3254-3263 (1994).
- Nelson, A. J., Hosier, S., Brady, W., Linsley, P. S. & Farr, A. G. *J. Immun.* **151**, 2453-2461 (1993).
- Mizuocho, T., Kasai, M., Kokuho, T., Kakiuchi, T. & Hirokawa, K. *J. exp. Med.* **175**, 1601-1605 (1992).
- Hoffman, M. W., Allison, J. & Miller, J. F. A. P. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2536-2530 (1992).
- Burky, L. C., et al. *J. Immun.* **151**, 3954-3960 (1993).
- Poirier, B., Lo, D., Reilly, C. R. & Kaye, J. *Immun.* **1**, 385-391 (1994).
- Inaba, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3038-3042 (1993).
- Ip, Y. T. et al. *Cell* **75**, 753-763 (1993).
- Scott, B. et al. *Immun.* **1**, 73-82 (1994).
- Lo, D. et al. *Eur. J. Immun.* **23**, 1693-1698 (1993).
- Arnold, B., Burgert, H.-G., Archibald, A. L. & Kvist, S. *Nucleic Acids Res.* **12**, 9473-9486 (1984).

ACKNOWLEDGEMENTS. We thank M. Carson for expert molecular biology advice and critical comments. These studies were supported by grants from the US National Institutes of Health to D. L.

†† To whom correspondence should be addressed.

TABLE 1 Substrate specificity of protein tyrosine kinases

Tyrosine kinases	-4	-3	-2	-1	Position 0	+1	+2	+3	+4
c-Fps/Fes EEEEIEIE	E (2.0) A * (1.9) D (1.5)	E (2.8) A (1.6)	E (2.7) A (1.5)	I (3.5) E (1.9) V (1.5)	Y	E (4.5) D (1.8)	E (2.7) D (2.7)	I (2.6) V (2.2) E (2.1)	E (2.7) I (1.8) V (1.6)
MT/c-Src EEIYGEFF	D (1.5)	E (2.5) D (1.9)	E (2.0) D (1.7) G (1.5)	I (3.0) V (2.4) L (1.5)	Y	G (2.2) E (2.1)	E (1.7)	F (2.9) I (1.5)	F (1.5)
v-Src EEEEIGEFD	E (1.7) D (1.5)	E (3.2) D (2.2)	E (2.2) D (2.0) G (1.6)	I (4.2) V (3.1) L (1.6)	Y	G (2.6) E (2.5) D (1.6)	E (2.5)	F (4.0) I (1.8) L (1.7) V (1.7)	D (2.2) E (2.1)
Lck XEXIYGVLF	X	E (1.5)	X	I (3.4) V (3.0) L (2.3) F (1.6)	Y	G (2.2) A (1.6)	V (1.6)	L (2.3) V (2.2) F (2.2) I (2.0)	F (1.7) L (1.7) V (1.6) I (1.5)
c-Abl AEVIYAAPF	A * (1.7)	X	V (1.5)	I (3.2) V (1.8) L (1.5)	Y	A (2.2)	A (1.6)	P (2.3) F (1.9)	F (1.5)
EGF receptor EEEEYFELV	E (1.6) D (1.5) R (1.5) A * (1.5)	E (2.9) D (1.8)	E (2.6) D (1.5)	E (2.5) D (2.1) I (1.6)	Y	F (2.1) V (1.9) I (1.7) E (1.5)	E (1.9) F (1.6) D (1.5)	L (1.7) I (1.7) F (1.6) V (1.5)	V (1.6)
PDGF receptor EEEEYVFIE	E (1.6) D (1.5)	E (2.0) D (1.5)	E (2.0) D (1.6)	E (2.2) D (1.6)	Y	V (2.0) E (1.7) I (1.5)	F (1.6)	I (1.9) V (1.7) F (1.5)	X
FGF receptor	A * (2.1)	E (1.6) A (1.5)	E (1.8)	E (1.7)	Y	F (1.9) V (1.5)	F (2.0)	L (1.8) F (1.8) M (1.7) I (1.6) V (1.6)	F (1.7)
Insulin receptor XEEYYMMMM	X	E (1.7)	E (1.8)	E (1.6) N (1.5) D (1.5)	Y	M (2.8) F (1.6)	M (2.2) F (1.6)	M (2.2) F (1.9) I (1.5) E (1.5)	M (1.9) F (1.9)

Values in parentheses indicate the relative selectivities for the amino acids⁴; amino acids with values less than 1.5 are omitted. Bold letters indicate amino acids that are strongly selected; X indicates no selectivity. The one-letter amino-acid code is used. The PTKs c-Fps/Fes (human), c-Abl (GST-fusion), v-Src, c-Src (chicken, co-expressed with polyoma middle T)²², Lck, EGF receptor²³, PDGF receptor, FGF receptor (cytosolic domain) and insulin receptor (cytosolic domain) were expressed in Sf9 cells using baculovirus. A protein-tyrosine kinase substrate library with the sequence⁴ Met-Ala-Xxx-Xxx-Xxx-Xxx-Tyr-Xxx-Xxx-Xxx-Xxx-Ala-Lys-Lys-Lys (where Xxx indicates any amino acids apart from Trp, Cys, Tyr, Ser or Thr) was presented to the purified PTKs. The peptide library (1 mg) was added to 300 µl of the protein kinase (0.5–10 µg) in a solution containing 100 µM ATP (with $\sim 6 \times 10^5$ c.p.m. [γ -³²P]ATP), 1 mM DTT, 10 mM MnCl₂, 50 mM NaCl and 50 mM Tris, pH 7.5. After 2.5 hours at 25 °C, the phosphopeptides were separated on DEAE-Sephacel and a ferric chelation column⁴. Typically, the ratio of product over enzyme was ~ 100 mol mol⁻¹ ($\sim 1\%$ of peptides phosphorylated). The phosphopeptide mixture was sequenced and data plotted as in ref. 4. Each kinase was evaluated at least twice; average values are shown.

* Partially due to lag from previous sequencing cycle.

-1 is consistent with a previous study of Abl substrate specificity⁹. In contrast, the receptor PTKs selected substrates with Glu at the -1 position and large hydrophobic amino acids at the +1 position. Every PTK examined preferred substrates with large hydrophobic amino acids at the +3 position, although the optimal amino acids differed.

In general, the sequences of known phosphorylation sites attributed to specific PTKs are in good agreement with the sequences predicted by the peptide library results^{10,11,25} (S. Blechner *et al.*, manuscript in preparation). For example, a Blast search of the SwissProtein database using the optimal peptide substrate predicted for the insulin receptor (EEEEMMMM) indicated that one of the best possible substrates of the insulin receptor from all known mammalian proteins is IRS-1, the major *in vivo* substrate of the insulin receptor. The Ret motifs¹² (hydrophobic-Tyr-X-X-Leu) on the T-cell receptor subunits are in good agreement with the optimal motif predicted for Lck (Table 1). The recently cloned phosphoprotein p130CAS is multiply phosphorylated on tyrosine residues in cells transformed by cytosolic PTKs¹³. Remarkably, most of the potential phosphorylation sites have β -branched amino acids (Val, Ile or Thr) at the Y1 position, in agreement with the optimal peptide substrates pre-

dicted for the cytosolic PTKs we studied (Table 1). Because the sequence I/VYXXP appears a dozen times in p130, our data predict that this protein would be an excellent substrate for the Abl PTK.

The optimal peptides predicted for the Src in Abl PTKs were synthesized and shown to be high-affinity and specific substrates for these enzymes (Fig. 1). The synthetic Src optimal peptide, AEEIYGEFEAKKKK, is an excellent substrate for Src ($K_m = 33$ µM, $V_{max} = 0.8$ µmol mg⁻¹ min⁻¹), but a poor substrate for Abl. Similarly, the Abl optimal peptide EAIYAAPFAKKK was favourably phosphorylated by Abl PTK ($K_m = 4$ µM, $V_{max} = 0.5$ µmol mg⁻¹ min⁻¹), but poorly phosphorylated by Src PTK (Fig. 1).

A comparison of the optimal peptide substrates of PTKs (Table 1) with the phosphopeptide motifs selected by individual SH2 domains^{3,14} reveals that individual PTKs preferentially phosphorylate sequences that bind to specific SH2 domains. The receptor PTKs phosphorylated peptides with the general motif Tyr-hydrophobic-X-hydrophobic, characteristic of peptides recognized by group III SH2 domains (such as p85, phospholipase C- γ , SHPTP2)³. In contrast, the SH2-containing PTKs preferentially phosphorylated peptides with the motif Tyr-hydrophilic-

hydrophilic-hydrophobic, recognized by group I SH2 domains³. Even more interesting is the finding that some of the SH2-containing PTKs selectively phosphorylate sites recognized by their own SH2 domains. For example, the optimal substrate for the Fps/Fes PTK contains the motif YEEI/V and the optimal phosphopeptide for the Fps/Fes SH2 domain is pYEEV/I (ref. 14). This selection is not due to the ability of the SH2 domain to protect the product from phosphatases or to present the substrate to the catalytic domain because similar motifs were found in experiments using kinase constructs lacking SH2 domains. The idea that cytosolic PTKs preferentially phosphorylate sites recognized by their own or related SH2 domains is supported by the data in Fig. 2a. The Abl SH2 domain bound a large fraction of peptides phosphorylated by the Abl PTK but a smaller fraction of peptides phosphorylated by the Lck PTK. In contrast, the Lck SH2 domain preferentially bound peptides phosphorylated by the Lck PTK. This result is consistent with the predictions of the library, because the c-Abl PTK (Table 1) and Abl SH2 domain³ both select peptides with the motif YXXP whereas the Lck PTK (Table 1) and Lck SH2 domain³ select for peptides with the motif YXXI/L/V. These results are also in agreement with the finding that proteins phosphorylated in *v-fps*-transformed cells preferentially bind to the Fps SH2 domain¹⁵.

Finally, evidence that catalytic site specificity is critical for fidelity in downstream signalling is provided by studies of the RET receptor PTK. A germline mutation affecting the catalytic core of this enzyme results in the autosomal dominantly inher-

ited cancer syndrome, multiple endocrine neoplasia type-2B (MEN2B)^{16,18}. The residue mutated is in a region of domain VIII that has been implicated in peptide substrate selectivity of protein Ser/Thr kinases^{4,19,20}. The wild-type RET protein has a methionine at residue 918 which is found in most other receptor PTKs. The mutation to threonine at this position makes this region more similar to SH2-containing PTKs and is expected to affect selectivity for amino acids at the +1 position⁴. The results in Fig. 2b demonstrate that the wild-type RET protein preferentially phosphorylates the optimal peptide substrate for the epidermal growth factor receptor, whereas the mutant RET

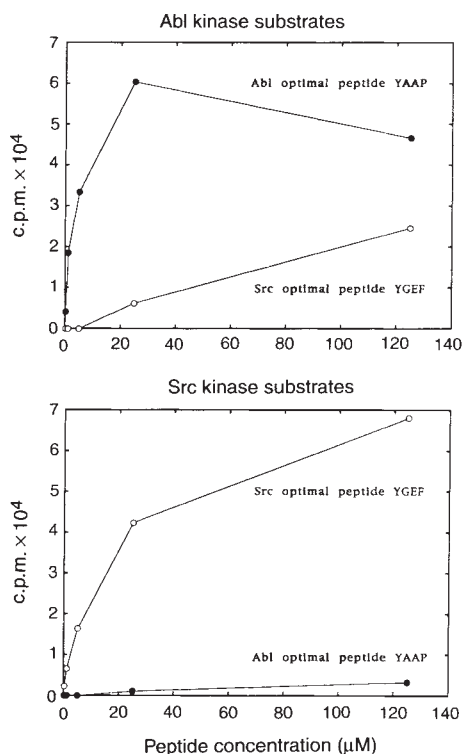


FIG. 1 Comparison of optimal substrates for Abl and Src kinases. Two peptides AEEIYGEFEAKKKK (Src optimal substrate) and EAIYAAPFAKKK (Abl optimal substrate) were phosphorylated by GST-SH2-Abl (lacking the SH3 domain) and GST-SH2-Src (constructed in the context of GST-SH2-Abl, where Abl SH2 and kinase domains were replaced by those of c-Src)²⁴. Different concentration of peptides were phosphorylated by GST-SH2-Abl or GST-SH2-Src in the kinase buffer (10 mM MnCl₂, 50 mM Tris, pH 7.5, 2 mM DTT), 10 μM ATP and 5 μCi [γ -³²P]ATP for 3 min at 25 °C. The amount of radioactivity incorporated was determined using the phosphocellulose assay.

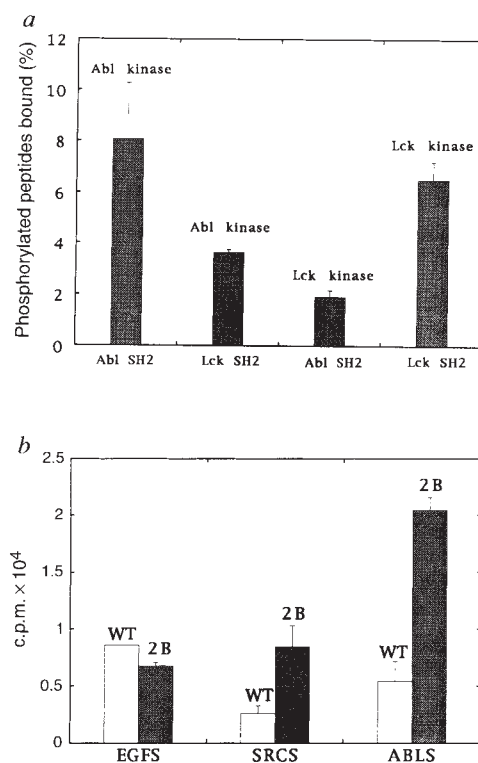


FIG. 2 Catalytic site specificities of PTKs are important for selective signalling. a, Overlapping specificities of SH2 domains and PTK domains were demonstrated using a peptide library. The degenerate peptide library MAXXXYXXXXAKKK was phosphorylated by purified Lck or GST-c-Abl with [γ -³²P]ATP such that a small fraction (~0.001%) of peptides was phosphorylated. Synthetic degenerate phosphopeptides³ (30 μM) were also added to the reaction buffer to prevent binding of the labelled peptides to endogenous SH2 domains in the kinases. The labelled phosphopeptide mixture was added to immobilized GST-SH2 domains (300 nM) with 1 μM synthetic phosphopeptides³. To correct for background binding, control experiments were performed in the same solution with 20 mM phenyl phosphate. After 2 hours incubation, the pellets were washed quickly with 1 ml ice-cold PBS (ref. 3) and the radioactivity associated with the pellets was measured. The plot shows percentages of total phosphorylated peptides binding to SH2 domains of c-Abl and Lck. Error bars indicate the standard errors (n=3). b, Baculovirus-expressed cytoplasmic domains of both wild-type RET (WT) or MEN2B mutant RET (2B) were assayed for protein kinase activity using three different peptide substrates. The RET proteins were immunoprecipitated from insect cells using phosphotyrosine-specific antibodies (4G10, UBI) and used to phosphorylate the EGF receptor optimal substrate (EGFS) AEEIYFELVAKKKK (10 μM), Src optimal substrate (SRCS) AEEIYGEFEAKKK (10 μM), or Abl optimal substrate (ABLS) EAIYAAPFAKKK (10 μM) in the presence of [γ -³²P]ATP. The amount of radioactivity incorporated was determined using the phosphocellulose assay. Immunoprecipitates from sf9 cells infected with virus lacking RET insert had no detectable kinase activity toward these peptides. The incorporated radioactivity was normalized such that RET WT and 2B had the same autophosphorylation activity. The error bars indicate the difference between two experiments.

preferentially phosphorylates optimal substrates for Src and Abl. These results are consistent with the models proposed here and suggest that phosphorylation of inappropriate substrates by the MEN2B RET PTK results in the stimulation of pathways normally mediated by SH2-containing PTKs, probably involving signalling proteins with group I SH2 domains.

In summary, we have shown that PTKs have distinct substrate specificities. Our findings suggest that PTKs and SH2 domains provide a double selection to maintain the fidelity of signalling events. In some cases, the ability of the SH2 domain of a PTK to bind a site phosphorylated by the catalytic domain of the same enzyme may allow processive phosphorylation of key substrates²¹ (B.J.M. *et al.*, manuscript submitted). Because an alteration in the catalytic-site specificity of RET results in cellular transformation, the catalytic-site specificities of PTKs must be important in maintaining fidelity in signalling. □

Received 14 November; accepted 20 December 1994.

1. Cantley, L. C. *et al.* *Cell* **64**, 281–302 (1991).
2. Pawson, T. & Schlessinger, J. *Curr. Biol.* **3**, 434–442 (1993).
3. Songyang, Z. *et al.* *Cell* **72**, 767–778 (1993).
4. Songyang, Z. *et al.* *Curr. Biol.* **4**, 973–982 (1994).

5. Patschinsky, T., Hunter, T., Esch, F. S., Cooper, J. A. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 973–977 (1982).
6. Hunter, T. *J. biol. Chem.* **257**, 4843–4848 (1982).
7. Cooper, J. A., Esch, F. S., Taylor, S. S. & Hunter, T. *J. biol. Chem.* **259**, 7835–7841 (1984).
8. Tinker, D. A., Cartron, J. L., McMurray, J. S. & Levin, V. A. *Anticancer Res.* **12**, 123–128 (1992).
9. Till, J. H., Annan, R. S., Cart, S. A. & Miller, W. T. *J. biol. Chem.* **269**, 7423–7428 (1994).
10. Pearson, R. B. & Kemp, B. E. *Meth. Enzym.* **200**, 63–81 (1991).
11. Valius, M. & Kazlauskas, A. *Cell* **73**, 321–334 (1993).
12. Reth, M. *Nature* **338**, 383–384 (1989).
13. Sakai, R. *et al.* *EMBO J.* **13**, 3748–3756 (1994).
14. Songyang, Z. *et al.* *Molec. cell. Biol.* **14**, 2777–2785 (1994).
15. Areces, L. B., Sbarba, P. D., Jucker, M., Stanley, E. R. & Feldman, R. A. *Molec. cell. Biol.* **14**, 4606–4615 (1994).
16. Hofstra, R. M. *et al.* *Nature* **367**, 375–376 (1994).
17. Eng, C. *et al.* *Hum. molec. Genet.* **3**, 237–241 (1994).
18. Carlson, K. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 1579–1583 (1994).
19. Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1988).
20. Knighton, D. R. *et al.* *Science* **253**, 414–420 (1991).
21. Dwyer, J. B., Skarvan, R. & Wang, J. Y. *J. Proc. natn. Acad. Sci. U.S.A.* (in the press).
22. Piwnicka-Worms, H., Williams, N. G., Cheng, S. H. & Roberts, T. M. *J. Virol.* **64**, 61–68 (1990).
23. Carraway III, K. L. & Cerione, R. A. *J. biol. Chem.* **268**, 23860–23867 (1993).
24. Mayer, B. J. & Baltimore, D. *Molec. cell. Biol.* **14**, 2883–2894 (1994).
25. Shoelson, S. E., Chatterjee, S., Chaudhuri, M. & White, M. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2027–2031 (1992).

ACKNOWLEDGEMENTS. We thank M. Berne (Tufts University) for peptide synthesis and sequencing. R.A.F. thanks J. Medin and K. Ozato for help in identifying recombinant c-Fps/Fes baculovirus; B.J.P. thanks T. Stonehouse for technical assistance. This work was supported by the Lucille P. Markey Charitable Trust and BASF BioResearch Co. Work in R.A.F.'s laboratory was supported by a grant from the National Cancer Institute.

Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 Å resolution

E. Y. Jones^{*†}, K. Harlos^{*†}, M. J. Bottomley[‡], R. C. Robinson^{*}, P. C. Driscoll^{††}, R. M. Edwards[§], J. M. Clements^{||}, T. J. Dudgeon[§] & D. I. Stuart^{*†}

^{*} Laboratory of Molecular Biophysics, [†] Oxford Centre for Molecular Sciences, [‡] Department of Biochemistry, South Parks Road, Oxford OX1 3QU, UK
[§] British Biotech Ltd, Watlington Road, Cowley, Oxford OX4 5LY, UK
^{||} NEURES Ltd, 4–10 The Quadrant, Barton Lane, Abingdon, Oxon OX14 3YS, UK

THE cell-surface glycoprotein vascular cell adhesion molecule-1 (VCAM-1; ref. 1) mediates intercellular adhesion² by specific binding to the integrin very-late antigen-4 (VLA-4, $\alpha_4\beta_1$; ref. 3). VCAM-1, with the intercellular adhesion molecules ICAM-1, ICAM-2, ICAM-3 and the mucosal vascular addressin MAdCAM-1, forms an integrin-binding subgroup of the immunoglobulin superfamily. In addition to their clinical relevance in inflammation, these molecules act as cellular receptors for viral and parasitic agents². The predominant form of VCAM-1 *in vivo* has an amino-terminal extracellular region comprising seven immunoglobulin-like domains. Functional studies^{4–7} have identified a conserved integrin-binding motif in domains 1 and 4, variants of which are present in the N-terminal domain of all members of the immunoglobulin superfamily subgroup. We report here the crystal structure of a VLA-4-binding fragment composed of the first two domains of VCAM-1. The integrin-binding motif (Q38IDSPL) is highly exposed and forms the N-terminal region of the loop between β -strands C and D of domain 1. This motif exhibits a distinctive conformation which we predict will be common to all the integrin-binding IgSF molecules. These, and additional data, map VLA-4 binding to the face of the CFG β -sheet, the surface previously identified⁸ as the site for intercellular adhesive interactions between members of the immunoglobulin superfamily.

The soluble fragment VCAM-d1,2 (human VCAM-1 residues

1 to 202, with no N-linked glycosylation sites, expressed in *Escherichia coli*) formed stable, highly ordered crystals which contained two molecules per crystallographic asymmetric unit⁹. Details of the crystallographic structure determination are reported in Table 1. The model comprises two copies of residues 1 to 199 plus 254 ordered water molecules. With the exception of two flexible loops (residues 139–147 and 177–186), the main-chain conformation and the majority of side-chain conformations are essentially identical for the two crystallographically independent copies. A representative portion of electron density is shown in Fig. 1a. The molecule (Fig. 1b) has overall dimensions of $20 \times 30 \times 80 \text{ Å}^3$ and in the relative spatial arrangement of the N-terminal domain 1, short linker region and domain 2, shows closest similarity to the extracellular region of the cell-adhesion molecule CD2 (ref. 8). In the full extracellular region of VCAM-1, residues 197 to 202 would fall within the second linker region and domain 3. Removed from this structural context, residues 200 to 202 are disordered for both copies of the VCAM-d1,2 crystal structure.

Details for structural superpositions of domain 1 (residues 1 to 89) with other immunoglobulin-like domains are given in Fig. 1c. The best structural correspondence is with telokin¹⁰ (r.m.s. distance of 1.25 Å for 76 Ca positions). This form of immunoglobulin domain has been termed an I-type domain and is predicted to occur in many cell-surface receptors¹¹. The VCAM-d1,2 structure provides the first details, to our knowledge, of this immunoglobulin motif in a cell-surface molecule. In general, domain 1 is a compact, rigid structure; the most prominent feature is formed by the CD loop. This loop is shorter than the equivalent loops in other immunoglobulin superfamily (IgSF) structures and follows a unique main-chain path (Fig. 1c). The second disulphide bridge (Cys 28 to Cys 75), linking the BC and FG loops (Fig. 1b), is a distinctive feature of this subgroup of the IgSF (Fig. 2). Domain 2 (residues 96–196) clearly conforms to the C2-type topology, but so far is the largest example of this class of immunoglobulin-like domain. Structural superpositions reveal closest similarity with domain 2 of CD2 (ref. 8) (r.m.s. distance of 1.2 Å for 65 Ca positions; Fig. 1d), the β -strands are of comparable length, but VCAM-1 domain 2 is elongated by inserts in several of the interstrand loop regions. In particular, the C'E and FG loops are unusually long and show high-temperature factors in both copies of the crystal structure, which implies inherent flexibility (Fig. 1b).

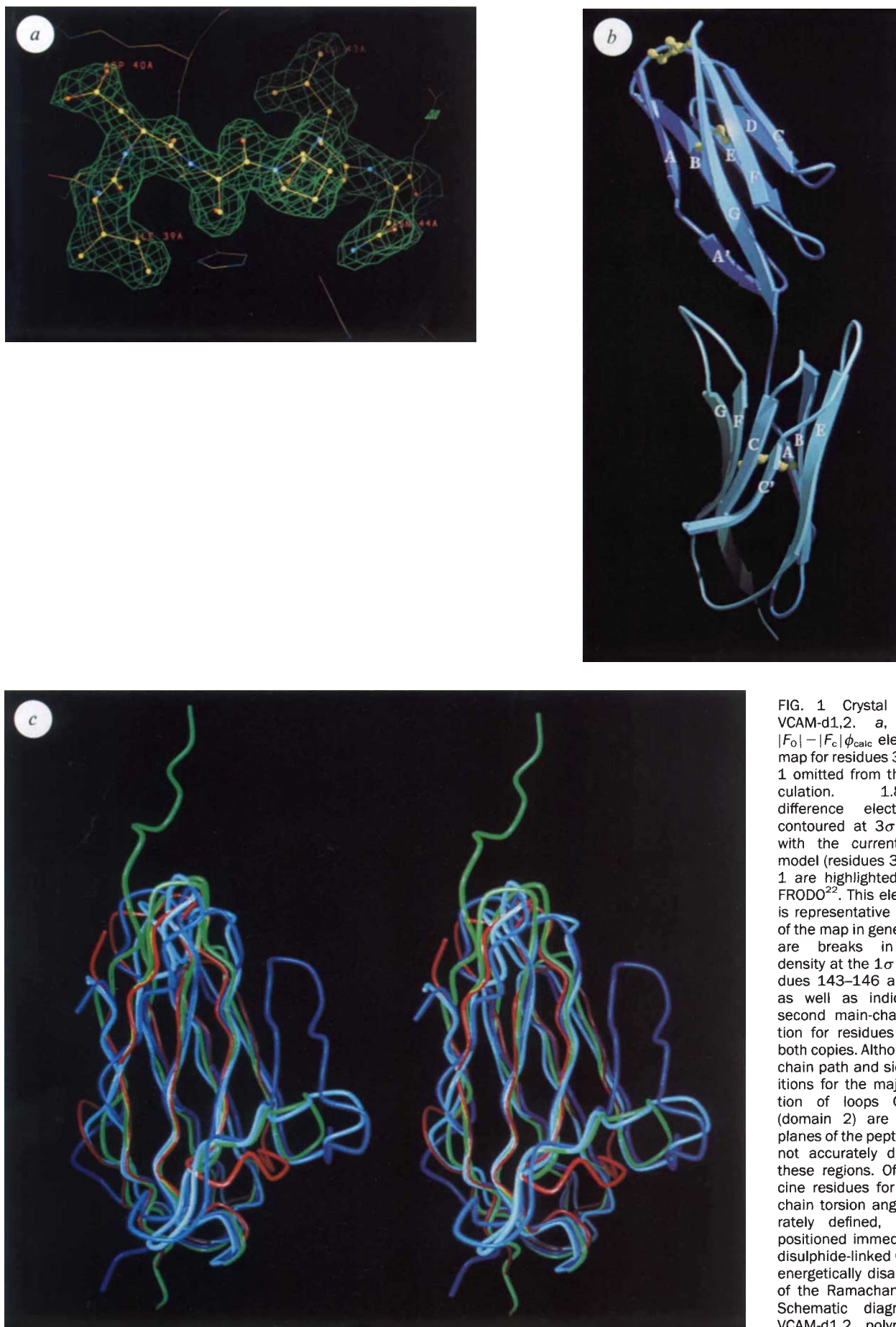
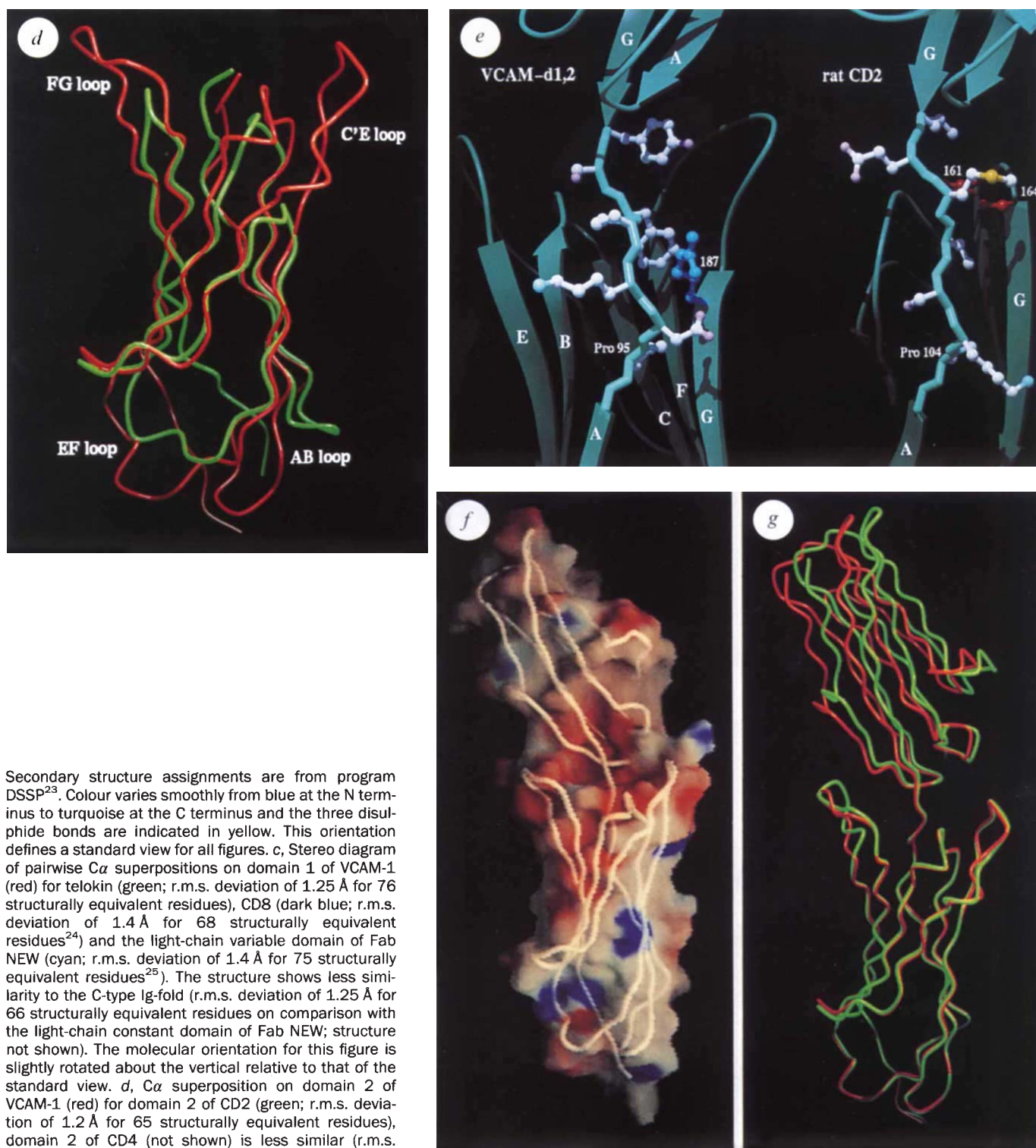


FIG. 1 Crystal structure of VCAM-d1,2. **a**, Portion of $|F_o| - |F_c| \phi_{\text{calc}}$ electron density map for residues 39–44 of copy 1 omitted from the phase calculation. 1.8 Å-resolution difference electron density contoured at 3σ is displayed with the current VCAM-d1,2 model (residues 39–44 of copy 1 are highlighted) in program FRODO²². This electron density is representative of the quality of the map in general, but there are breaks in main-chain density at the 1σ level for residues 143–146 and 180–184, as well as indications of a second main-chain conformation for residues 179–185 in both copies. Although the main-chain path and side-chain positions for the major conformation of loops CE' and FG (domain 2) are defined, the planes of the peptide bonds are not accurately determined in these regions. Of the non-glycine residues for which main-chain torsion angles are accurately defined, only Glu 76, positioned immediately after a disulphide-linked Cys, lies in an energetically disallowed region of the Ramachandran plot. **b**, Schematic diagram of the VCAM-d1,2 polypeptide fold. ►



Secondary structure assignments are from program DSSP²³. Colour varies smoothly from blue at the N terminus to turquoise at the C terminus and the three disulphide bonds are indicated in yellow. This orientation defines a standard view for all figures. *c*, Stereo diagram of pairwise $C\alpha$ superpositions on domain 1 of VCAM-1 (red) for telokin (green; r.m.s. deviation of 1.25 Å for 76 structurally equivalent residues), CD8 (dark blue; r.m.s. deviation of 1.4 Å for 68 structurally equivalent residues²⁴) and the light-chain variable domain of Fab NEW (cyan; r.m.s. deviation of 1.4 Å for 75 structurally equivalent residues²⁵). The structure shows less similarity to the C-type Ig-fold (r.m.s. deviation of 1.25 Å for 66 structurally equivalent residues on comparison with the light-chain constant domain of Fab NEW; structure not shown). The molecular orientation for this figure is slightly rotated about the vertical relative to that of the standard view. *d*, $C\alpha$ superposition on domain 2 of VCAM-1 (red) for domain 2 of CD2 (green; r.m.s. deviation of 1.2 Å for 65 structurally equivalent residues), domain 2 of CD4 (not shown) is less similar (r.m.s. deviation of 1.7 Å for 58 structurally equivalent residues) owing to a distinctive broadening of its β -barrel. *e*, Comparison of the VCAM-d1,2 (residues 89–95) and CD2 (residues 98–104) linker regions. The molecular orientation is rotated from the default by some 180° about the vertical. A proline residue marks the transition to strand A of domain 2 in both molecules. Side chains that mediate hydrogen bonds to the linker are also shown: Arg 187 in VCAM-d1,2 (blue), Asn 161 and Ser 164 in CD2 (red). Subtle variations of main-chain torsion angles in the linker region cause changes of up to 20° in the relative orientation of domains 1 and 2 in comparisons of the human and two rat CD2 crystal structures¹⁶. The relative orientation of the domains in VCAM-d1,2 conforms most closely (35° difference) to that seen in human CD2. *f*, Electrostatic potential surface of VCAM-d1,2. Regions of positive

potential are shown in blue, and regions of negative potential are shown in red. The opposite face of the molecule to that displayed is relatively neutral. *g*, Diagram of the copy-1 VCAM-d1,2 structure (red) compared to the copy-2 structure (green). Equivalent $C\alpha$ atoms are superposed for domain 2. Panels *b–e* and *g* were drawn using programs MOLSCRIPT (ref. 26, with modifications by R. Esnouf) and RASTER3D (E. A. Merritt and M. E. P. Murphy); structural superpositions were made using the program SHP²⁷; electrostatic analysis and surface display for *f* used program GRASP²⁸.

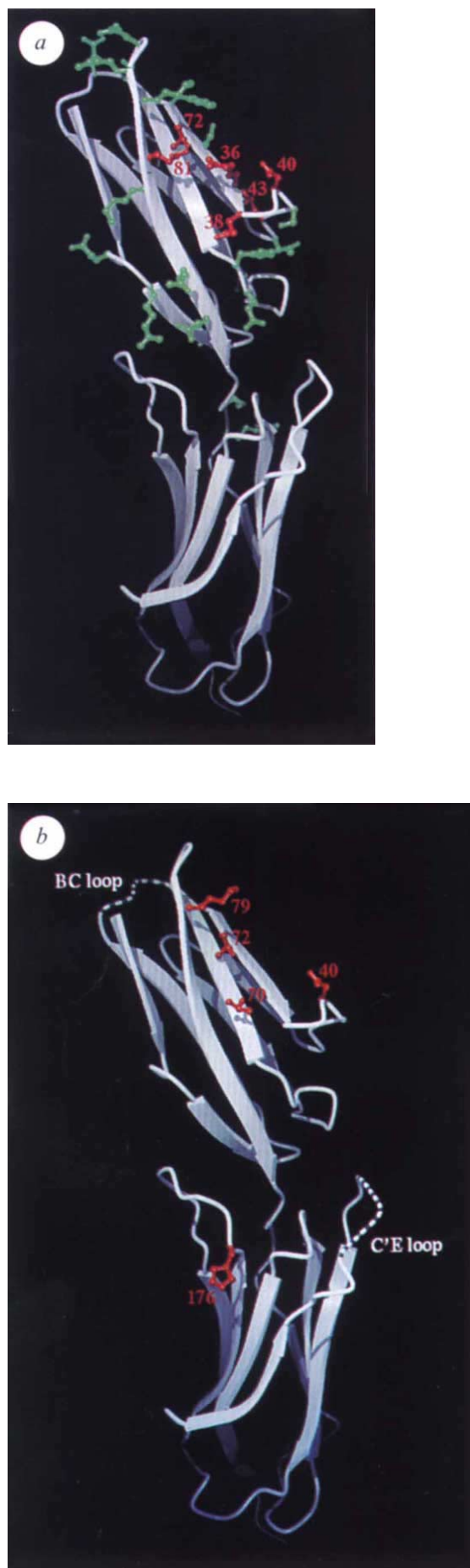


FIG. 2 Sequence alignments, based on structural considerations, for human VCAM-d1,2, VCAM-d4,5 and the N-terminal two domains of human MAdCAM, ICAM-1, ICAM-2 and ICAM-3. In the consensus sequence, upper case represents a totally conserved amino acid, lower case a strongly conserved amino acid, and ! conservation of chemical type. Secondary structural assignments and solvent-accessible surface areas for VCAM-d1,2 were calculated in program DSSP. Asterisk indicates residues with less than $10 \text{ \AA}^2$ of solvent-accessible area that are not involved in the domain interface; '&' indicates residues that contribute $>10 \text{ \AA}^2$ of buried surface area to the VCAM domain 1/domain 2 interface; and Δ indicates hydrophobic residues exposed in VCAM-d1,2 that are likely to contribute to the domain 2/domain 3 interface in the full-length molecule. Exposed VCAM-d1,2 residues at which site-directed mutagenesis⁴⁻⁷ does not affect integrin binding are indicated by a plus sign, and residues at which effects occur are indicated by a minus sign. The integrin-binding motif is shaded. Several distinctive structural features of VCAM-1 domain 1 appear to be conserved within the family: the proline (Pro 7 in VCAM-1) and consequent strand A to A' switch, the additional disulphide bridge at the top of the β -barrel, and the N-terminal portion of the CD loop.

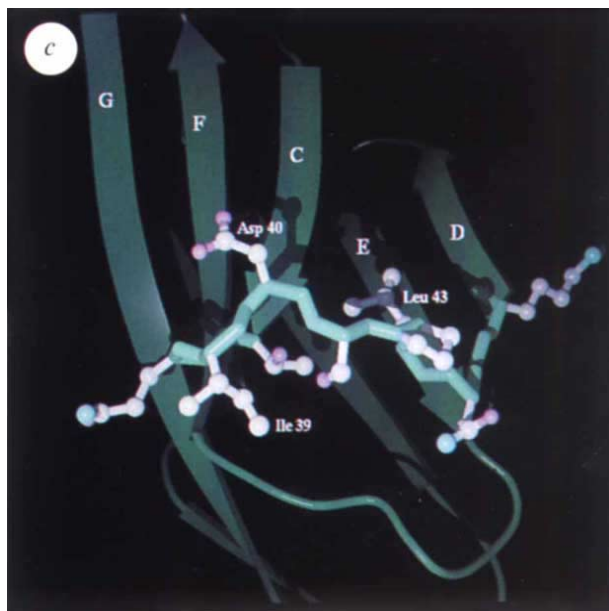


FIG. 3 Integrin-binding sites mapped to the VCAM-d1,2 structure. *a*, Schematic diagram of VCAM-d1,2, with those side chains of exposed residues (greater than $10 \text{ \AA}^2$ of solvent-accessible area) for which mutagenesis⁴⁻⁷ does affect adhesive function indicated in red, and those at which function is not affected indicated in green. Mutated residues are listed in Fig. 2; residues 36, 38, 40, 43, 46, 72 and 81 have been implicated in adhesion, but the role of three of these (38, 43 and 46) in direct integrin binding should be qualified. A Gln-to-Ser mutation at residue 38, at the start of the CD loop, has been reported to affect function⁵, although an alanine mutation had no significant effect⁴. The loss-of-function mutation of the semi-buried Leu 43 (ref. 5) may arise from perturbation of the CD loop conformation as this side chain packs against the hydrophobic core of the domain. Lysine 46, situated at the start of strand D, appears to be an outlier in the cluster of functionally significant residues, but as the functional data on this residue is for mouse VCAM-1 (ref. 6), it may contribute to the stability of the CD and/or EF loop in a species-specific manner. *b*, Schematic diagram of VCAM-d1,2 with exposed side chains (residues 40, 70, 72, 79 and 176), whose structural equivalents in ICAM-1 or ICAM-3 are implicated^{14,15} in integrin LFA-1 binding, indicated in red. Loop regions predicted to be truncated in the ICAMs are indicated by black stripes. *c*, VCAM-1 integrin-binding motif. The CD loop in domain 1 of VCAM-1. All components of this figure were drawn using programs MOLSCRIPT (with modifications by R. Esnouf) and RASTER3D.

	<--A-->	<--A-->	<--B-->	<--C-->	<-->
	10	20	30	40	
VCAM12	FKIETT-PESRYLAQIGDSVSLTCTGCESPFFSWRTGSPIN--GKV				
VCAM45	----ftveis--pppriaaigidsvmtscvmgcespsfswrtgspis--gkv				
MAdCAM	----qsfqvnppesevavamtalgitcsmcdgcvahvhwrtgslg--svq				
ICAM1	----qtqsvs--pskvilprgg--slvtvtcsts--cdapklldatpdkkell				
ICAM2	sdekvfvehvr--pkklavepvg--slevncstt--cnpqevvgstetnk--ill				
ICAM3	--qefllrve--pqpvlvsagg--slfvncstd--cpssekialerisk--elv				
Consensus:	P.....G.S.....CS.....L.....				
Buried:	*.....&&&&.....*.....*.....*.....				
Function:	*.....*.....*.....*.....*.....*.....				

	D-->	<--E-->	<--F-->	<--G-->	<--A-->
	50	60	70	80	90
VCAM12	TNEGTTSTLTMPVSVFNGEHSYLCATATCESRKLEKGIQVEIYSPFKDPEIHLGSP				
VCAM45	rsegtntstlslpsvfenehsylctvtcghkklekigqvlysfprpdeiemsgg				
MAdCAM	tlpg--saillsvrgmls--dtgtpvcvsgcgsrfqhsvkilyafpdqlvvspefl				
ICAM1	lpgnarkvyslnvq--edsqpmcysncpdqgqstaktfltvwtpervelapls				
ICAM2	degaqkwhylvsnis--hdtvlqchftcsgkqesmsnsvsygpprvlcltqpt				
ICAM3	asgmgaafnlsnvt--gnsrilcsyvcngsqitgssnitvyglpervelaplp				
Consensus:	!.....!.....!.....!.....!.....!.....!.....				
Buried:	*.....*.....*.....*.....*.....*.....*.....				
Function:	*.....*.....*.....*.....*.....*.....*.....				

	<--B-->	<--C-->	<--C'-->	
	110	120	130	140
VCAM12	L-EAGKIPITVKCSVADVPFD--RLEIDLLKGDHLMKSGEFL-----DADR			
VCAM45	l-vngssvtvscvskpvypld--rlelellkgetileniefl-----dtdm			
MAdCAM	v--pgqdvqvstahniwpadpnsllsfalllgeqrlegaqalepeqeieqaeg			
ICAM1	wpgvgnknlrlrcqvqegapra--nlvtvllrgekelkrepaug-----			
ICAM2	lvavgsfctiercvptvepld--sltlflfrgnetlhyetfgka-----a			
ICAM3	wpgvgnknlrlrcqvqegsprt--slvtvllrweeelsrqpave-----			
Consensus:	!...G...!...C.V...!...P...!...L...!...L...!...L...!...P			
Buried:	*...Δ...*...*...*...Δ...*...*...*...&&&&.....*...Δ...*			

	<--E-->	<--F-->	<--G-->	
	150	160	170	180
VCAM12	KSLETKSLVETFTFVIEDIGKTLVCRALKHI--DEMDSVPTVROAVKELQVYISP			
VCAM45	kslenkslemftiptiedtgkalcvaklhi--ddmefepkqrqstqtllyvnap			
MAdCAM	tpflrmtqwrslpslgtppapalhcvmtql--pk--lvltlrkelpvlsqstsp			
ICAM1	-epaevtttvlvrr--dhnganfscrteldlrrpqqlelfentsapyqlgtfvlp			
ICAM2	papgeatcatfnstadredghrnfscslavldlmsrggnifhkhspakmleiyepv			
ICAM3	-epaevtatvlasr--ddhgafpfcrteldmqpgglglfvntsapqlrtfvlp			
Consensus:	!...!...!...!...!...!...!...!...!...!...!...!...!...!...P			
Buried:	&&&.....*...Δ...*...Δ...*...*...*...&&&&.....*...Δ...*			

The interdomain linker (residues 90–95) is closely associated with the domain-2 β -barrel and follows a comparable path to the equivalent region in CD2 (Fig. 1e). Electrostatic calculations reveal that it lies at the centre of an extensive region of negative electrostatic potential (Fig. 1f). Residues from strand A' and the EF loop of domain 1, the linker and the BC, C'E and FG loops of domain 2, are involved in the domain interface (detailed in Fig. 2). Although the amount of buried area (850 Å²) is similar to that between domains 1 and 2 of CD4 (880 Å²; refs 12, 13), the presence of a distinct linker region, rather than a direct continuation of the domain-1 strand G into the domain-2 strand A, means that the relative arrangement and orientation of the two domains is more closely comparable with that of CD2 (buried area, 400 Å²). The increase in buried area arises from the extended C'E and FG loops of the VCAM-1 domain 2 and, by analogy with CD4, might have been expected to rigidify the structure. These additional contacts do not lock the relative domain orientation, however. Comparison of the two independent copies of the molecule in the crystallographic asymmetric unit reveals a relative flexion of 6° (Fig. 2g), with the FG loop behaving as a flexible buffer, reorienting in response to shifts in domain-1 position.

In Fig. 3a, functional data^{4–7} from mutagenesis of VCAM-1 are presented in the context of the structure. Surface residues implicated in VLA-4 (integrin) binding are clustered on the CFG face of the domain-1 β -barrel and, in particular, on the structurally distinctive CD loop. Asp 40, which is highly exposed on the CD loop, has been highlighted as a key functional residue^{4–7}. Indeed, the N-terminal portion of the CD loop (Q38IDSPL) appears to be vital in directly mediating integrin binding. In the full VCAM-1 extracellular region, a second copy of the integrin-binding motif (Q326IDSPL) would occupy a similarly exposed position on the CD loop of domain 4, thus providing a second distinct binding site. Any direct contribution of VCAM-1

domain 2 in integrin binding has not been tested. Both the flexible C'E and FG loops in domain 2 are abnormally long in comparison to other C2-type domain structures (Fig. 1d), and both loops, in particular the C'E loop, are close to the integrin-binding area on domain 1. Residues from these loops contribute to the concentration of negative charge which forms the most striking feature of the VCAM-d1,2 electrostatic potential surface (Fig. 1f). This region is thus a prime candidate for mutagenic studies.

A sequence alignment for the integrin-binding subgroup of the IgSF, based on the structure of VCAM-d1,2, is presented in Fig. 2. It is probable that the VCAM-d1,2 structural scaffold recurs in each of these molecules. The positions of functionally important residues in ICAM-1 (ref. 14) and ICAM-3 (ref. 15), which unambiguously map to surface positions in the VCAM-d1,2 structure, are displayed in Fig. 3b. The CFG face of the domain 1 β -barrel is again implicated in integrin binding. The nature and function of the N-terminal portion of the CD loop appears to be identical across the whole family; indeed, the interchangeability of the ICAM and VCAM forms of this motif has been demonstrated experimentally⁴. Thus all these leukocyte integrin-binding molecules mediate their adhesive interactions through a prominently exposed hydrophobic side chain (isoleucine or leucine), followed by a negatively charged aspartic or glutamic acid.

Mutation of Asp 166 in ICAM-3 (ref. 15) (equivalent to His 176 in VCAM-d1,2; Fig. 2), which maps to the start of the domain-2 FG loop, implicates this domain in integrin binding. The loop regions in domain 2 show substantial differences in length between VCAM-1, MAdCAM and the ICAMs (Fig. 2). The C'E loop, in particular, is extended in VCAM-1, and even more prominently in MAdCAM, compared to the ICAMs. In both VCAM-1 and MAdCAM, these inserts bear arrays of negatively charged aspartic and glutamic acids. As both these molecules can compete for binding to the same integrin, $\alpha_4\beta_7$, elements of integrin-binding specificity within the IgSF molecules may reside in the C'E loop of domain 2. The length of the BC loop is the most obvious gross structural difference in domain 1 by which to sub-classify these molecules. Again, VCAM-1 and MAdCAM fall within the same structural subgroup, with the ICAMs all showing a markedly truncated BC loop region. These structural subgroups parallel an obvious sequence variation between the integrins, with ICAMs, but not VCAM-1 or MAdCAM, binding to integrins that contain an additional inserted domain (I domain).

Crystal structures for the extracellular region of rat⁸ and human¹⁶ CD2 provided a model for intercellular adhesive interactions between IgSF molecules. Such interactions are mediated through the CC'C'FG face of the N-terminal domains. The crystal structure of VCAM-d1,2 now indicates that the CFG face of the domain 1 β -barrel may serve as the platform for intercellular adhesive interactions by IgSF molecules, irrespective of the nature of their ligand. The N-terminal portion of the CD loop provides a structurally distinctive motif for the IgSF/integrin class of adhesion interaction. This structural motif may occur more generally in integrin-binding molecules as a related sequence motif (Leu-Asp-Val) has been shown to be essential for VLA-4 mediated cell adhesion by the alternatively spliced type-III connecting segment domain of fibronectin^{7,17}. The tripeptide Arg-Gly-Asp plays a key role in mediating binding for a range of molecules that interact with the non-leukocyte integrins. Arg-Gly-Asp peptide structures are potent competitors of integrin binding and one of these inhibits the VCAM-1/VLA-4 interaction¹⁸. However, the conformation of the CD loop in VCAM-d1,2 bears little similarity to the Arg-Gly-Asp structure seen in various integrin-binding proteins (see for example, refs 19 and 20) other than in the presence of a highly exposed, negatively charged side chain located in a flexible loop region. Such side chains may contribute to adhesion by completing a divalent cation-binding site in the integrin. Our results will help in the

TABLE 1 Statistics for crystallographic structure determination

Crystals	$P2_12_12_1$		$a = 52.7 \text{ \AA}; b = 66.5 \text{ \AA}; c = 113.2 \text{ \AA}$			45% Solvent content		
Diffraction data								
Data set	$d_{\min} \text{ \AA}$	No. of obs. No. of crystals	Unique reflection % complete	R_{merge} (all data)	MFID to 3 \AA	No. of sites	Phasing power (3 \AA)	
Native SRS	2.0	175,783 (3)	23,615 (94)	10.5	—	—	—	
Native KEK	1.8	142,904 (1)	31,634 (84)	6.7	—	—	—	
Native ALL	1.8	—	32,077 (85)	8.7	—	—	—	
Se-Met KEK	2.5	123,785 (2)	13,935 (97)	7.7	6.7	4	0.5	
EMP LAB	3.0	49,665 (1)	7,662 (91)	9.9	14.5	2	0.9	
UAc SRS	2.6	74,174 (3)	12,213 (98)	5.9	21.1	2	0.3	
Refinement: general statistics								
Resolution (\AA)	No. of reflections	No. of atoms	R_{cryst} (%)	Δ_{bonds} (\AA)	Δ_{angles} ($^{\circ}$)	$\Delta_{\text{B}_{\text{bond}}}$ \AA ²	$\Delta_{\text{B}_{\text{ang}}}$ \AA ²	
15.0–1.8	32,030	3,362	20.3	0.015	2.01	2.97	4.71	
Refinement: non-crystallographic symmetry								
Domain	Δ_{pos}	copy 1 vs 2	Δ_{B}	copy 1 vs 2	Mean B	copy 1	Mean B	copy 2
1		0.22		3.8		18.5		21.5
2		0.25		2.2		31.7		29.8

Production and crystallization of VCAM-d1,2 and selenomethionyl (Se-Met) VCAM-d1,2 have been described^{9,21}. Diffraction data sets were collected at ambient temperatures for: in-house (LAB) using an 18 cm diameter MAR-research imaging plate system (J. Hendrix and A. Lentfer) on a Rigaku RU200 rotating anode ($\lambda = 1.54 \text{ \AA}$); SERC Synchrotron Radiation Source, Daresbury (SRS) using a MAR-research imaging plate system (18 cm diameter) at station 9.5 ($\lambda = 0.99 \text{ \AA}$ or $0.90 \text{ \AA}$); the Photon Factory, Japan (KEK) on Fuji BASIII imaging plates using the Weissenberg method at station BL6A2 ($\lambda = 0.97 \text{ \AA}$). Data sets were auto-indexed (KEK data using 1° oscillation images recorded for the purpose) then integrated with program DENZO (Z. Otwinowski). Heavy-atom soak conditions comprised 10 mM ethyl mercury phosphate for 22 h (EMP); 10 mM uranyl acetate for 26 h (UAc). Native and derivative (including Se-Met) data were scaled and isomorphous, and anomalous difference Patterson maps were calculated and solved using standard methods. Heavy-atom refinement and phasing were performed using the CCP4 version (CCP4: SERC Collaborative Computing Project no. 4, Daresbury Laboratory, Warrington, UK, 1979) of program MLPHARE (Z. Otwinowski). The figure of merit for multiple isomorphous replacement (MIR, including information from anomalous scattering) phases to $3.0 \text{ \AA}$ resolution was 0.44 on all reflections. ($R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$; MFID, mean fractional isomorphous difference; phasing power, mean value of the heavy-atom structure factor amplitudes divided by the residual lack-of closure error.) MIR phases were modified by solvent flattening (program GAP; J. Grimes and D. Stuart) and typical Ig-like domains manually positioned in the $3.0 \text{ \AA}$ resolution solvent-flattened map to define the molecular envelopes. Non-crystallographic symmetry operators were then refined by density correlation and the phases improved by iterations of solvent flattening and real-space averaging using GAP (initial electron density correlation coefficient, 0.46; final electron density correlation coefficient, 0.79). A first model of 180 residues (including two regions of polyalanine) was built for one copy of VCAM-d1,2 in the $3.0 \text{ \AA}$ resolution two-fold averaged map, using the program FRODO²² and the second copy generated by non-crystallographic symmetry. Rigid-body refinement of domains in the program X-PLOR (A. T. Brunger, X-PLOR Manual, Version 3.1. Yale University) revealed a slight variation in the relative orientations of the N- and C-terminal domains for copies 1 and 2 of VCAM-d1,2, tight non-crystallographic restraints were therefore applied separately to domains 1 and 2. Refinement by conjugate gradient minimization, restrained temperature factor, and simulated annealing progressively incorporated all data 8.0 – $2.0 \text{ \AA}$. Several cycles of rebuilding and refinement improved the model and extended the resolution to $1.8 \text{ \AA}$ (rebuilding on copy 1 was initially at $3.0 \text{ \AA}$ resolution, using averaged maps based on combined MIR and model phases, and latterly at $1.8 \text{ \AA}$ resolution using $2|F_o| - |F_c| \phi_{\text{calc}}$ maps; copy 2 was regenerated by applying non-crystallographic symmetry to the rebuilt copy 1). A bulk solvent correction allowed all measured data from 15.0 to $1.8 \text{ \AA}$ to be used. In the final stages of the refinement, ordered water molecules were added to the model, the non-crystallographic symmetry restraints released and models for both copies rebuilt. (Δ_{bonds} and Δ_{angles} , root-mean-square deviations from ideal bond lengths and angles; $\Delta_{B_{\text{bond}}}$ and $\Delta_{B_{\text{ang}}}$, root-mean-square deviations in B-factors between bond and bond-angle related atoms; Δ_{pos} and Δ_B , root-mean-square deviations in position and B-factor for main-chain atoms between copy 1 and copy 2 excluding residues 130–134, 139–147, 177–186 and 194–199; mean B, average temperature factor for main-chain atoms).

design of therapeutic approaches targeting the integrin-binding subgroup of the IgSF, but a full understanding of the structural determinants for this form of cell adhesion must await more information on the integrins. □

Received 3 November; accepted 29 December 1994.

- Osborn, L. et al. *Cell* **59**, 1203–1211 (1989).
- Springer, T. A. *Nature* **346**, 425–434 (1990).
- Elices, M. J. et al. *Cell* **60**, 577–584 (1990).
- Osborn, L. et al. *J. Cell. Biol.* **124**, 601–608 (1994).
- Vonderheide, R. H., Tedder, T. F., Springer, T. A. & Staunton, D. E. *J. Cell. Biol.* **125**, 215–222 (1994).
- Renz, M. E. et al. *J. Cell Biol.* **125**, 1395–1406 (1994).
- Clements, J. M. et al. *J. Cell Science* **107**, 2127–2135 (1994).
- Jones, E. Y., Davis, S. J., Williams, A. F., Harlos, K. & Stuart, D. I. *Nature* **360**, 232–239 (1992).
- Bottomley, M. J. et al. *J. molec. Biol.* **244**, 464–468 (1994).
- Holden, H. M., Ito, M., Hartshorne, D. J. & Rayment, I. *J. molec. Biol.* **227**, 840–851 (1992).
- Harpaz, Y. & Chothia, C. *J. molec. Biol.* **238**, 528–539 (1994).
- Wang, J. et al. *Nature* **348**, 411–418 (1990).
- Ryu, S.-E. et al. *Nature* **348**, 419–426 (1990).

- Staunton, D. E., Dustin, M. L., Erickson, H. P. & Springer, T. A. *Cell* **61**, 243–254 (1990).
- Holness, C. L. et al. *J. biol. Chem.* (in the press).
- Bodian, D. L., Jones, E. Y., Harlos, K., Stuart, D. I. & Davis, S. J. *Structure* **2**, 755–766 (1994).
- Komoriya, A. et al. *J. biol. Chem.* **266**, 15075–15079 (1991).
- Cardarelli, P. M. et al. *J. biol. Chem.* **269**, 18668–18673 (1994).
- Logan, D. et al. *Nature* **362**, 566–568 (1993).
- Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P. *Science* **258**, 987–991 (1992).
- Dudgeon, T. J. et al. *Eur. J. Biochem.* **226**, 517–523 (1994).
- Jones, A. *Meth. Enzym.* **115**, 157–171 (1985).
- Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).
- Leahy, D. J., Axel, R. & Hendrickson, W. A. *Cell* **68**, 1145–1162 (1992).
- Saul, F. A., Amzel, L. M. & Poljak, R. J. *J. biol. Chem.* **253**, 585–597 (1978).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 945–949 (1991).
- Stuart, D. I., Levine, M., Muirhead, H. & Stammers, D. K. *J. molec. Biol.* **134**, 109–142 (1979).
- Nicholls, A., Sharp, K. A. & Honig, B. *Proteins* **11**, 281–296 (1991).

ACKNOWLEDGEMENTS. We thank the staff at the SRS, Daresbury, and at the Photon Factory, Japan, for help with X-ray data collection; R. Bryan and R. Esnouf for computing facilities and programs; R. Esnouf with D. Stammers and J.-S. Ren for assistance at the Photon Factory; S. Lee for help in the preparation of figures; D. Simmons for discussion; and J.-H. Wang for exchange of information on crystal characteristics. The Oxford Centre for Molecular Sciences is supported by the BBSRC and MRC; E.Y.J. and P.C.D. are supported by Royal Society Research Fellowships; and R.C.R. by an MRC research studentship. Atomic coordinates for VCAM-d1,2 have been deposited with the Protein Data Bank, Brookhaven National Laboratory, USA.